DOCUMENT RESUME

ED 066 492 TM 001 976

| | |
|---|---|
| AUTHOR | Brown, Les; And Others |
| TITLE | Evaluation for Individualized Instruction. An Operational Guide for Teacher Workshops. |
| INSTITUTION | Institute for Educational Research, Downers Grove, Ill. |
| PUB DATE | 72 |
| NOTE | 87p. |
| AVAILABLE FROM | Institute for Educational Research, 1400 West Maple Avenue, Downers Grove, Illinois 60515 ($4.00) |
| EDRS PRICE | MF-$0.65 HC-$3.29 |
| DESCRIPTORS | *Behavioral Objectives; Curriculum Development; Evaluation Techniques; Guides; *Individualized Instruction; Inservice Teacher Education; *Item Banks; Program Evaluation; Teacher Role; *Teacher Workshops; Teaching Skills |
| IDENTIFIERS | ESEA Title III; *Evaluation for Individualized Instruction Project |

ABSTRACT

The workshops associated with the Evaluation for Individualized Instruction Project are planned to provide teachers an opportunity to improve their skills in construction and development of evaluation procedures. This is based on the premise that one successful initial approach to the improvement and individualization of instruction is through the improvement of teacher-made tests. Presented in this operational guide for teacher workshops are eleven Instructional Sessions, they are: One--Educational Objectives; Two--Defining Global Objectives; Three--Behavioral Objectives; Four--Hierarchy of Learning Behaviors; Five--Purposes of Achievement Testing; Six--Techniques of Measurement; Seven--Multiple-Choice Items; Eight--Basic Statistical Measures; Nine--Test Analysis Interpretation of Computer Output; Ten--Validity and Reliability; and Eleven--The Liaison Role of the Workshop Participant. A list of reference materials is included. (LS)

ED 066492

ED 001 926

# EVALUATION FOR INDIVIDUALIZED INSTRUCTION

A Title III ESEA project
administered by
Downers Grove, Illinois
School District 99

INSTITUTE FOR EDUCATIONAL RESEARCH

1400 West Maple Avenue
Downers Grove, Illinois 60515
Phone: 312-971-2040

1972 EDITION

## AN OPERATIONAL GUIDE FOR TEACHER WORKSHOPS

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

# EVALUATION FOR INDIVIDUALIZED INSTRUCTION

1400 West Maple Avenue
Downers Grove, Illinois 60515
(312) 971-2040

## AN OPERATIONAL GUIDE FOR TEACHER WORKSHOPS

A Title III ESEA Project

Marcus Lieberman, Director
Administered by Downers Grove, Illinois,
School District No. 99

This guide represents the most recent revision
of an evolving curriculum developed by

Les Brown, Georgia Brooks,
Patricia J. Cocks and Mildred E. Kersh

CONSULTANTS

Norman V. Bowers PhD.
John Wick PhD.

1972 EDITION

TABLE OF CONTENTS


| | Page |
|---|---|
| Introduction | 1 |
| Instructional Sessions: | |
| One - - - - - - - Educational Objectives | 5 |
| Two - - - - - - Defining Global Objectives | 11 |
| Three - - - - - Behavioral Objectives | 13 |
| Four - - - - - - Hierarchy of Learning Behaviors | 25 |
| Five - - - - - - Purposes of Achievement Testing | 29 |
| Six - - - - - - - Techniques of Measurement | 31 |
| Seven - - - - - Multiple-Choice Items | 47 |
| Eight - - - - - - Basic Statistical Measures | 55 |
| Nine - - - - - - Test Analysis Interpretation of Computer Output | 69 |
| Ten - - - - - - - Validity and Reliability | 81 |
| Eleven - - - - - The Liaison Role of the Workshop Participant | 87 |
| Reference Materials | 90 |

## I. NEED FOR THE PROJECT

The need for this project arises from a number of closely associated ideas. These ideas are all related to important instructional tasks undertaken consistently and frequently by classroom teachers.

One such idea is that it is unfortunate that instruction is limited by tests that are narrow in scope and improperly conceived. A second idea is a conviction that teachers can build suitable tests that measure much more than a low-level recall of information. The third idea is that teachers can utilize test data in the process of individualizing instruction. The fourth idea is that students can profit from receiving feedback about their learning. The fifth idea is that teachers, and indeed all others interested in the educational process, need to know how the pupils are performing more often than once each year.

The workshops associated with the Evaluation for Individualized Instruction Project are planned to provide teachers an opportunity to improve their skills in construction and development of evaluation procedures. This is based on the premise that one successful initial approach to the improvement and individualization of instruction is through the improvement of teacher-made tests.

These project workshops should produce (1) a group of teachers who are sophisticated in the construction of evaluative instruments and in the techniques for using test data to improve the individualization of instruction; (2) an available pool of teacher-made test items that are appropriate for measuring both information and the broader aspects of education; (3) the recommendations to students of remedial procedures to be followed if tested skills and understandings have not been mastered.

To attain these ends, immediate as well as long-range objectives for the workshop have been developed. These are stated in terms of the terminal behavior of the participants.

## A. IMMEDIATE OBJECTIVES

The immediate instructional objectives of the workshop are those which can be evaluated as each workshop progresses and immediately after termination of the workshop.

During the course of the workshop and as a result of instruction received, the teacher can

1. distinguish between properly and improperly stated instructional objectives.
2. write instructional objectives stated in terms of student behavior and applied to a particular area of content.
3. criticize multiple-choice items in terms of the anatomy, grammar, and logic of the form as well as the plausibility, clearness and accuracy of the content.
4. write well-formed multiple-choice test items which test behaviors critical to attainment of important instructional objectives.
5. demonstrate understanding that the objectives of instruction should include attainment of cognitive behaviors above the level of rote recall by writing multiple-choice items that test, not only recall, but also higher levels of cognitive behavior such as "comprehension, understanding, application, analysis, synthesis."[1]
6. participate in the compilation of a complete list of content specified for areas of application and employed by the member school districts.

## B. LONG–TERM OBJECTIVES

The long-term instructional objectives of the workshop are those which depend on implementation by teachers in their districts over an extended length of time.

As a result of participation in the Evaluation Project workshop teachers can

1. recall that the three major purposes of the Evaluation Project are
   a. to provide teachers with improved skills in the construction, development, and use of classroom evaluation procedures.
   b. to demonstrate that such improvement in teacher skill will lead to the improvement of programs of individualized instruction to promote mastery of essential objectives.
   c. to provide a model for the use of continuous evaluation information in both improving instruction and in giving information to communities regarding the progress of students in their schools.
2. demonstrate their understanding of the three major purposes of the Evaluation Project by
   a. answering questions about the project.
   b. demonstrating and teaching the skills of objective item writing.

[1]Taxonomy of Educational Objectives, Benjamin S. Bloom, et al.

c. explaining uses of the item pool and of the information that can be gained by use of data from item pool tests.

d. using the item pool themselves to construct tests in their own subject area.

e. using the results of item pool tests to analyze their own instructional programs and pupil needs.

f. revising their own instruction in the light of information gained through statistical analysis of test results so that

(1) classroom instruction is more effective.

(2) individual pupil needs are better understood and are more adequately met.

3. demonstrate their knowledge and understanding of their "liaison" role in the Evaluation Project by

a. reporting in writing to their school districts--summarizing their workshop activities and detailing ways in which they believe skills they have gained in the workshop can be used to help improve evaluation procedures in their districts.

b. assisting in changing evaluation procedures where requested by their district administration.

c. requesting assistance from the Evaluation Project Staff in planning and carrying out a program of information and education.

d. using the techniques for assisting innovation learned in the Evaluation Project workshop.

4. demonstrate their appreciation of the values of the project by

a. participating with positive enthusiasm in the workshop projects.

b. sharing ideas and offering suggestions during conferences and in general sessions of the workshop.

c. attending voluntarily, special meetings and offering their services outside their own districts to help other participants where such assistance is requested and the districts approve the cooperative effort.

d. providing a personal model of proper use of evaluation tools and techniques in their own subject areas by drawing upon the item pool for the construction of tests and by using the results to evaluate on-going instruction and to meet individual pupil need in the classroom.

## PLAN OF OPERATION: EVALUATION PROJECT WORKSHOPS

The plan of operation of the workshops intersperses instructional sessions, led by one or more of the project staff, with work sessions in which the participants write and edit instructional objectives and multiple-choice items. Simultaneously, the staff maintains contact with former workshop participants by supplying them with materials and observing evaluation techniques employed in their classes.

A. **INSTRUCTIONAL SESSIONS:** Teachers receive instruction in (1) the basic concepts of evaluation and measurement, (2) the writing of instructional objectives, (3) the writing of well-formed multiple-choice test items, (4) the construction of balanced classroom tests, (5) the analysis of test results, and (6) the application of test analysis to classroom instruction. Especially stressed are the benefits to individual students resulting from the use of test data to plan remediation, enrichment and acceleration.

   The instructional content of the workshops is carefully planned but the approach is informal. Teachers are encouraged to question the information given, to seek clarification and to explore problems in relation to the concepts being taught.

   This portion of the workshop is evaluated by means of a multiple-choice post test, and by sampling from the objectives and items written by teachers throughout the workshop sessions. The objectives and items used to evaluate instruction and teacher progress are checked against a list of errors similar to that used by the teachers to edit their items.

B. **WORK SESSIONS:** Teachers write behaviorally stated instructional objectives and test items. Staff associates work closely with participants, helping them to analyze and edit the form and content of the objectives and items they write.

C. **LIAISON:** The role of the workshop teacher in assisting change in the school is explored. Teachers learn that even after completing the workshop, they may expect the continuing support of the project staff in (1) providing up-to-date information on the project, (2) helping them to fulfill the expectations of their school and district with regard to the objectives of the project.

   Workshop teachers are requested to report in writing to their district, summarizing their work as a participant in the project. They are asked to describe the ways they think the skills they have acquired by participating in the Evaluation Project can be used to produce better learning in their schools.

   A meeting is held near the end of each workshop. This meeting brings together previous and current workshop participants. In this way, all participants are kept up-to-date on the project and the meeting also provides a valuable opportunity for sharing ideas.

**Instructional Session One:** **Educational Objectives**

**A. OBJECTIVES**

The participant

1. can demonstrate knowledge of educational objectives by identifying specific changes in their form over the past century.
2. can comprehend the relationship between objectives, evaluation and learning activities by giving specific examples of each and describing their interdependence.
3. can state the purpose of objectives in evaluating the effectiveness of innovative programs.
4. can demonstrate knowledge of the sources of educational objectives as specified by Tyler by naming the particular source considered in an illustrative example.

**B. INSTRUCTIONAL PROCEDURES:**

1. Lecture on the history of educational objectives
2. Discussion of the purpose of educational objectives
3. Lecture on the relationships between objectives, learning activities, and evaluation

**C. PARTICIPANT'S PREPARATION:**

1. Part I of Talmage's Instructional Objectives
2. Tyler's Basic Principles of Curriculum and Instruction, pp. 3-28

**D. INSTRUCTIONAL PROCEDURE:** Discussion of the sources of educational objectives

**E. MATERIALS:**

1. Instructional Cycle: chart on the relationships between objectives, learning activities, and evaluation.
2. Part I of Talmage's Instructional Objectives
3. Tyler's Basic Principles of Curriculum and Instruction, pp. 3-28

**F. EVALUATION:** Proficiency Quiz No. 1

STUDY GUIDE – SOURCES OF CURRICULUM

A. Miss Raven is beginning her career as a mathematics teacher in an economically disadvantaged area school. Approximately five percent of the school's graduates continue their education. She has an excellent series of units on calculus that she designed while student teaching in a wealthy suburban community. Since she had such a fine response to these materials while student teaching, she decides to use them in her present situation. The children do very poorly with them. Miss Raven is convinced that they could master the material if her students would "only study." Which of the following sources of curriculum was probably being ignored?

1) subject matter specialists
2) children's needs
3) desires of the local community

B. Give a reason for your choice.

## EVOLUTION OF EDUCATIONAL OBJECTIVES IN THE U. S. SINCE 1900.

**1900 - W. W. I**

The mind was thought to have certain faculties, such as memory and reason, which could be trained or disciplined by proper exercise.

Objectives were stated in terms of the mental faculty to be trained. The study of Latin was considered especially valuable as its mastery required memory, while its grammatical structure provided exercise in orderly reasoning.

**1918 - 1925**

Learning was perceived as the acquisition of highly specific patterns of behavior. Each task was viewed as a stimulus to which the student was to learn the appropriate response.

Examples:

1. Able to add 2 + 3
2. Able to add 2 + 4
3. Able to add 2 + 5
4. Able to spell believe
5. Able to spell receive

Difficulties: Number of objectives for each curricular area was phenomenal. One course in arithmetic listed 3,000 objectives.

This view of objectives declined after 1925 for two reasons:

1. Teacher could not keep in mind so many objectives.
2. The approach denied the development of generalized behavior

**1930's - Present**

The Eight-Year Study sharply articulated the idea that learning can best be defined as a change in pupil behavior and therefore, objectives should be directed towards obtaining this change.

Examples:

1. The Development of Effective Ways of Thinking
2. The Development of Effective Work Habits and Skills
3. Increased Sensitivity to Social Problems
4. Greater Awareness of Esthetic Experiences
5. An Appreciation of Man's Relation to his Physical Environment

Difficulties: Objectives of this type are usually too general to provide direction to the classroom teacher and yet they are still found frequently in curriculum guides.

Mid-Century Committee on Outcomes in Elementary Education

"Behavior = Action by a Pupil"

Examples:

1. The pupil knows the customs of India.
2. The pupil can use the library for obtaining further materials
3. The pupils enjoy participating in group activities.

The work of James Popham, Robert Mager and others

"Behavior = Observable Action by a Pupil"

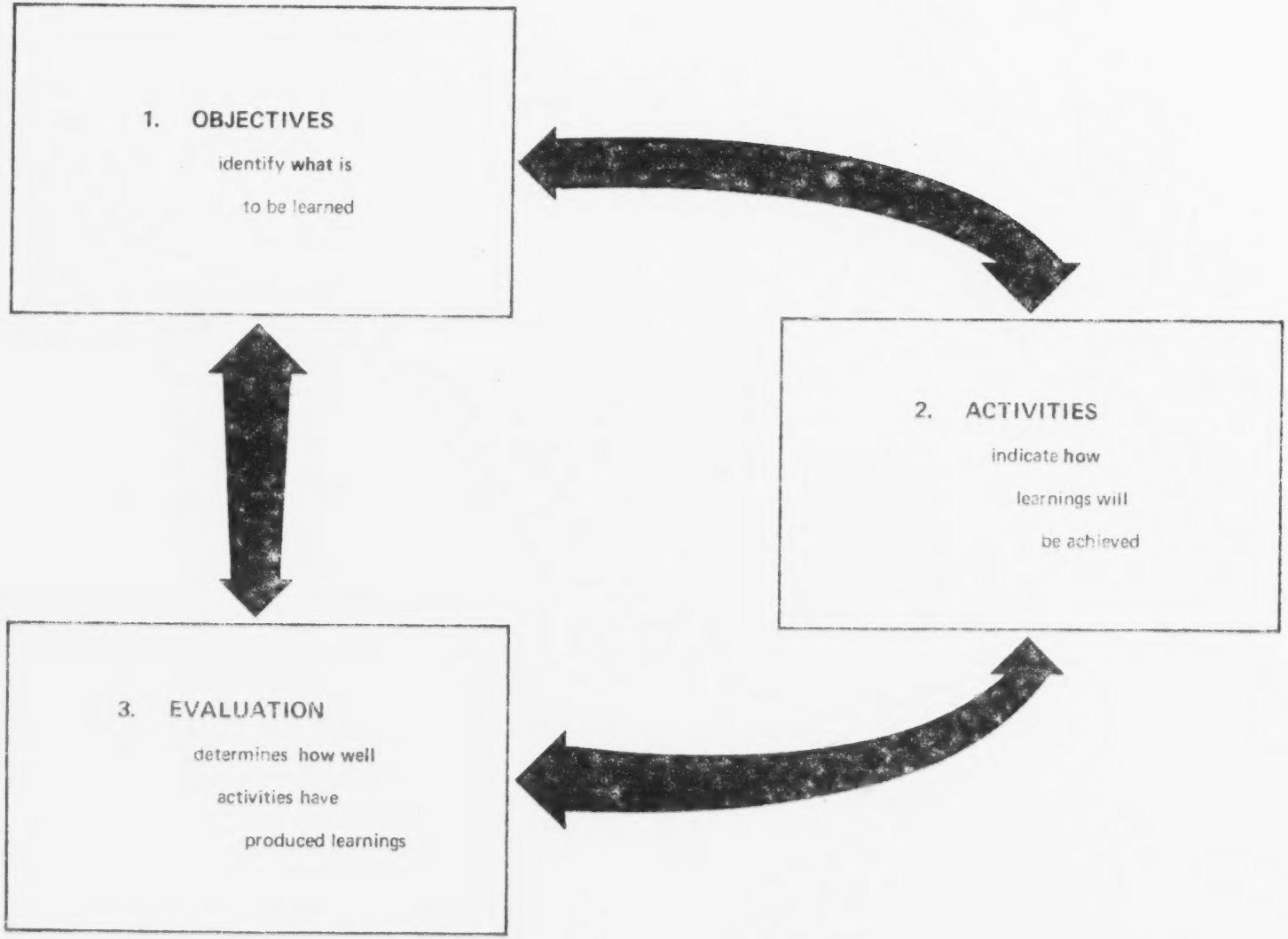
Examples:

1. The student will circle the names of the five most recent presidents of the United States when given a list of ten past presidents.

2. The student will solve four of five multiplication problems with the product less than twenty-five.

3. Given a set of arguments and conclusions, the student will correctly select 80% of those whose conclusions are warranted.

## INSTRUCTIONAL CYCLE



Instructional Session Two:

## Defining Global Objectives

**A. OBJECTIVES**

The participant

1. can describe the purpose of stating educational objectives in various forms, e.g., public relations, global, instructional.
2. can, from a list of objectives, indicate which are globally stated and which are behaviorally stated.
3. can rewrite global objectives into a behaviorally observable form.
4. will listen openmindedly to the philosophy of stating objectives behaviorally as evidenced by remarks made in a discussion on the topic.
5. will attempt to write behaviorally stated objectives that help to measure the attainment of a global objective.
6. will react in a rational, knowledgeable manner to objections to the philosophy of behavioral objectives as evidenced by rebuttals to stated arguments.
7. will accept change in student behavior as a measure of attainment of objectives rather than subjective opinions, as evidenced by the participant creating problematic situations and making judgments about proficiency based on performance.

**B. INSTRUCTIONAL PROCEDURES:**

1. Participants will complete a List of Selected Global and Behavioral Objectives.
2. Discussion of the necessity of defining objectives in terms of pupil behavior if they are to provide direction to the teacher in guiding learning activities.

**C. PARTICIPANT'S PREPARATION:** Part II of Talmage's Instructional Objectives

**D. INSTRUCTIONAL PROCEDURES:**

1. Provide practice in defining global objectives in terms of student behavior.
2. Provide practice in rewriting global objectives in behavioral terms.
3. Group critique of efforts to rewrite global objectives.

**E. MATERIALS:**

1. Part II of Talmage's Instructional Objectives
2. List of Selected Global and Behavioral Objectives

**F. EVALUATION:** Proficiency Quiz No. 2.

## A LIST OF SELECTED GLOBAL AND BEHAVIORAL OBJECTIVES

Place an X before the objectives which you feel would provide valuable guidance to you in planning learning activities to realize them

1. _____ The student will be able to correctly thread a sewing machine.
2. _____ The student will develop a sense of the cultural unity of man.
3. _____ The student will list and describe the themes of Shelley's poems.
4. _____ The child will develop interest in music.
5. _____ The child will demonstrate an interest in Mark Twain's works by reading and reporting on novels or short stories not required in class.
6. _____ The student will understand fundamental concepts of physics.
7. _____ The pupil will recognize the incorrect usage of capital letters when he employs rules given in class.
8. _____ The student will learn about collective-bargaining.
9. _____ The student will be able to write an essay in which he contrasts the arguments for having a democracy or a dictatorship.
10. _____ The student will be able to set up and conduct an experiment to determine the density of rocks.
11. _____ The pupil will demonstrate a real appreciation of the contributions made by the late Senator Everett Dirksen.
12. _____ The student will learn the parts of speech.
13. _____ The pupil will be able to list the bones in the human foot.
14. _____ The child will analyze an argument by identifying assumptions, conclusions, and emotionally laden statements.

Instructional Session Three:

## Behavioral Objectives

### A. OBJECTIVES

The participant

1. can state at least three reasons for restating global objectives into a behavioral form.
2. can indicate from a list of verbs, which are directly observable and which must be observed indirectly.
3. can indicate from a list of objectives, which are stated behaviorally and which are not.
4. can indicate, from a list of objectives, which are stated in terms of the student and which are stated in terms of the course or teacher.
5. can indicate, from a list of objectives, those that are explicitly referent to a concept or content and which are not.
6. can rewrite imperfectly stated objectives into acceptable, behavioral form.
7. can write behavioral objectives in his own grade and subject area, using the global objectives from curriculum materials as a guide.
8. can state the three domains of student behavior (cognitive, affective and psychomotor) and give examples of each in behavioral form.
9. can classify given behavioral objectives into cognitive, affective and psychomotor categories.
10. will place behavioral objectives in their proper perspective with respect to evaluation and learning activities as evidenced by remarks made in discussions.
11. will explain to others the philosophy and methodology of behavioral objectives upon completion of the workshop and returning to his district.
12. will encourage others to write behavioral objectives to evaluate pupil progress and innovative programs in his district as evidenced by the forming of program committees and requests for assistance from the project staff.

### B. INSTRUCTIONAL PROCEDURES

1. View Popham's filmstrip, Instructional Objectives.
2. Present Popham's paper, Instructional Objectives as a written record of the filmstrip presentation.
3. Discussion of the anatomy of a properly stated behavioral objective.
4. Prepare and discuss lists of directly and indirectly observable verbs.
5. Provide practice in rewriting behavioral objectives.
6. Provide practice in writing behavioral objectives in the participant's own subject matter areas.
7. Group critique of behavioral objectives written by participants.

C. **PARTICIPANT'S PREPARATION:** Mager's Preparing Instructional Objectives.

D. **INSTRUCTIONAL PROCEDURES:** Discussion of Mager's approach to writing behavioral objectives

E. **MATERIALS:**

1. Filmstrip, Instructional Objectives and answer sheets
2. List of directly and indirectly observable verbs to be prepared by participants
3. Anatomy of a Behavioral Objective
4. Mager's Preparing Instructional Objectives

F. **EVALUATION:** Proficiency Quiz No. 3

## INSTRUCTIONAL OBJECTIVES* W. James Popham

This program deals with the manner in which those involved in teaching or the development of instructional materials should specify their goals. The topic is one of paramount importance for the prospective product developer.

As you go through the program you will be asked to inspect certain frames, and then to respond on the answer sheet which has been provided. Use a 5 x 7 card as a mask so that you do not inadvertently see the right answer which is presented below the three asterisks. Sometimes you will see three asterisks in the middle of the page as this:

* * *

When you do, use your card to cover the area below the asterisks until you have responded on the answer sheet, then check for the correct answer below the asterisks and proceed with the program. Now begin the formal part of the program.

Pupils go to school to get an education. And most of us would agree on the basic aims of the American schools. We want to produce good citizens, competent workers, and sound thinkers. However, most educators realize that such "basic" aims are usually stated in terms so broad as to defy precise interpretation. If teachers wish to really implement these objectives they must have a more definite idea of what the objective means.

Teachers have always worried about the importance of instructional objectives, yet the kinds of objectives which teachers have endorsed, usually made little difference in the nature of the instructional program. The principal reason for this is that these objectives were stated in terms too broad and too ambiguous to allow anyone to agree upon what the objective meant.

In the next few frames, you will see some examples of actual objectives appearing in courses of study or lesson plans in public school use today.

**"At the end of the course the student will have developed qualities of rational thought and good citizenship."**

For this first objective try to decide what the objective means and what kind of an educational program would be built in order to achieve the objective. Do you think others viewing this objective would reach a similar conclusion regarding its meaning and an appropriate educational program? Answer yes or no by circling your answer on the answer sheet alongside Number 1.

The answer is **No**, for the objective is broad and permits a variety of interpretations. How about this objective?

**"The student will gain an appreciation of the importance of Western European literature."**

What kind of educational program would be planned to achieve it and, more importantly, how could one tell whether or not the objective had been achieved? Do you think most people would agree on the meaning of this objective? Circle your answer alongside Number 2.

* * *

*With permission, the material in this program is reproduced with minor modifications from the script of the copyrighted filmstrip-audiotape program Educational Objectives, by W. James Popham, Vimcet Associates, Los Angeles, California, 1966.

The answer should be No, for although this objective limits the content to Western European literature, it is still extremely vague. Objectives such as these are not much help in planning instruction. Suppose a school staff wishes to determine systematically whether it has been successful in achieving this kind of objective. What types of examinations or what sort of observations of the students might help in judging whether the objective has been achieved? The same question can be asked of the following objectives--how can one tell whether they have been successfully achieved?

The student will grasp the significance of the environment in which he lives.

Students will become familiar with the basic concepts of biology.

Do you think most educators would agree on the kinds of tests which would reflect satisfactory achievement of these goals? Answer on your answer sheet by Number 3.

* * *

Again the answer is No. There would be great disagreement over the way to measure attainment of these objectives.

Unless objectives unambiguously communicate what the educator intends to accomplish, they are of little instructional value. In fact, one might think of a continuum in which educational objectives become more useful as they become less ambiguous.

But why are ambiguous objectives of little value to the teacher? Since educators have been writing vague objectives for so many years, maybe they are worthwhile. Undoubtedly, they have some public relations value in that the schools can communicate general aims to the public with such broad statements. But in the classroom, broad, vague objectives are next to worthless because a teacher really can't make any instructional plans from them. Objectives should allow the teacher to know where he is going, that is; they should permit the teacher to formulate instructional plans for himself and his students so that the objective will be achieved. But if the objective is so vaguely stated that a number of interpretations is possible, is the objective useful? Yet, for no good reason, vague objectives are popular with too many classroom teachers.

For instance, many teachers and curriculum workers describe objectives in terms of concepts topics, or generalizations to be treated in the course.

The course will cover:

1. The Industrial Revolution.
2. The Civil War.
3. The Reconstruction Period.

These objectives are examples of such "content coverage" goals. But further consideration will reveal that the teacher does not merely wish to cover these topics--rather, the teacher expects that coverage to help the pupils become educated. Statements such as these objectives, then, really miss the point of education.

Other teachers and curriculum workers may state objectives in terms of what the teacher is to do during the course.

During this unit the teacher will:

1. Analyze the causes of World War II.
2. Discuss the background of the U.N.
3. Point out the current perils to peace.

Yet, even here we realize that the teacher should really be interested in how his activities influence his pupils. If the only real objective in a history class was for the teacher to analyze the causes of World War II, no pupils need be present. The teacher could "analyze" in an empty room. Such statements, therefore, fail to provide the teacher with any explicit guidance about ways the instruction should affect the students.

Some teachers also state objectives in extremely broad terms, with no specific reference.

The students will:

1. Develop appreciations.
2. Increase his interests.
3. Develop conceptual thinking.

These objectives, because of their nonspecific nature give little help to the systematic planning of instruction. Almost anything could be done.

What then is a good instructional objective? How should instructional goals be stated so that they are of value to the teacher in selecting learning activities and in subsequently evaluating to see whether the objectives have been accomplished? Good instructional objectives must be stated in terms of student behavior. Yes, meaningful and valuable instructional objectives must be described by stating how the student behaves, or will be able to behave, after instruction. The more specifically the pupil behavior can be defined the better. For when the teacher's goal is to change the student's observable behavior, a way of judging whether the objective has been achieved is provided--namely, to observe whether the behavior change has occurred. But what kind of "behavior"? How should we describe the way a pupil will behave after instruction? Is this a satisfactory description?

The pupil will understand.

When a teacher says that a student will "understand," precisely what does he mean? For instance, think about the kind of evidence you could use to demonstrate that this next objective had been achieved.

The student will understand the meaning of the Monroe Doctrine.

Or this one?

The student will really understand the meaning of the Monroe Doctrine.

It is true that most people have a somewhat similar conception of what the word "understand" means, but if you ask them to be more specific, for example, to describe the kind of student behavior, even test behavior, that reflects "understanding," you will find tremendous differences in their interpretation of this term.

For instance, some would think the student understood the meaning of the Monroe Doctrine when he could write out a description of it from memory; or maybe when he could answer a series of true-false questions about its background; or perhaps only when he could properly identify instances where the Doctrine was violated.

It is clear that terms such as "understanding," and "knowledge," and "insight," allow considerable latitude with respect to their interpretation. Even though terms such as understanding refer to the student, it is next to impossible to tell what they mean unless one further specifies what type of student behavior signifies understanding.

In the next frame, does phrase A or phrase B allow for fewer interpretations?

A. To select answers correctly.
B. To realize fully.

Circle the correct answer on your sheet by Number 4.

* * *

Phrase A is clearer because it describes an observable form of behavior.

Phrase B describes a rather vague, internal type of response which, because it is unobservable, might be interpreted in many ways.

In the next frame, pick the two phrases which identify a form of observable student behavior.

At the end of the unit the student will be able:

A. To differentiate.
B. To enjoy.
C. To comprehend.
D. To construct.

Circle the appropriate two letters on your answer sheet by Number 5.

* * *

A and D are the correct answers, for "differentiating" and "constructing" identify behavior which is observable, while "enjoying" and "comprehending" refer to internal, unobservable responses of the individual.

In this next frame select the two behavioral phrases and circle your answer sheet by Number 6 accordingly.

A. To think
B. To repair
C. To answer
D. To appreciate

* * *

B and C are the correct answers, for "repairing" and "answering" are behavioral, whereas "thinking" and "appreciation" are not at all specific.

In the next frame, is objective A or B less ambiguous?

A. The pupil will learn his multiplication tables.
B. The pupil will solve 4 or 5 multiplication problems.

Circle the correct answer by Number 7.

* * *

Objective B is clearer because **solving** is an identifiable form of pupil behavior. The pupil's responses on a test would serve as a behavioral index of his solutions, whereas **learning** is one of those non-behavioral words which is open to many different interpretations.

Which of these objectives is the more specific?

A. The student will know the novels of Faulkner.
B. The student will describe three literary movements of the Romantic Period.

Circle the best objective by Number 8 on your answer sheet.

* * *

Objective B is the more specific since it identifies the student behavior which is sought--namely, written description. Objective A hinges on the interpretation of "know" and this could be something as simple as merely listing the names of Faulkner's novels or as difficult as writing critical analysis.

In the next three frames select the behaviorally stated objectives by circling the appropriate letter on your answer sheet.

A. The student will be able to match the names of world leaders with their country.
B. The concept of world unity will be treated.

For this frame answer by Number 9.

* * *

In the preceding frame, the behaviorally stated objective is A. Choose the behaviorally stated objective for this next frame and answer by Number 10.

A. The student will have familiarity with the major stylistic schools in English verse.
B. The student will be able to write a 6 line poem which incorporates iambic pentameter.

* * *

Objective B is the correct answer for the frame. Now choose the best answer for this frame and answer by Number 11 on your answer sheet.

A. The student will really want to become a good teacher.
B. When given two objectives, the student will be able to correctly identify the one which is properly stated.

* * *

In this frame, objective B is correct, for while it is important to want to become a good teacher, it is difficult to determine whether or not a student really wants to. On the other hand, identifying properly stated objectives, as in choice B, is observable behavior--and incidently, will contribute much toward one's becoming a good teacher.

Now that you have had considerable practice in identifying behaviorally stated objectives, here is a more difficult task--the writing of such objectives.

Is the objective in this frame stated behaviorally?

The pupil will know the names of the five most recent presidents of the United States.

Circle the correct answer on your answer sheet by Number 12.

* * *

Of course not! For while this is a rather limited objective, the verb "know" is ambiguous because there are a number of observable student behaviors which might be used to reflect such knowledge. Now, on

your answer sheet, change this objective so that it is behavioral. That is, write out an improved objective which there is no ambiguity with respect to the behavior which reflects knowledge of the names of the five most recent presidents of the United States.

* * *

Look at your objective and decide whether there is an observable student behavior described. If the behavior of the student can be observed while he demonstrates this knowledge or if the product of his knowledge is observable, then the objective is acceptable. For example,

**The students will correctly name aloud the five most recent presidents of the United States.**

This objective describes a student behavior that can be observed as it takes place, and thus it would qualify as a behavioral objective.

This next objective describes a product, the list of circled names, which could be observed.

**The student will circle the names of the five most recent U. S. presidents from a list of all of our presidents.**

This too is a behavioral objective. Most objectives involving student test performance are of this type.

Other acceptable modifications of the objective might be the following:

**The student will list in writing the names of the five most recent presidents of the United States.**

**When presented with 5 multiple choice questions, each presenting one of the five most recent presidents and three distractors, the student will choose the correct answers.**

If your modified objective was similar to these, you have prepared an acceptable objective. Incidentally, if your behavioral objective requires student performance on a test, it is necessary to carefully describe the type of test that will be used. If possible, examples of the test items should be given, for it may be tempting to avoid the difficult task of explicitly describing student behavior by hiding behind such vague objectives involving test performance.

The next objective is definitely non-behavioral. Change it on your answer sheet so that it is behavioral.

**The student will gain a working knowledge of the RCA tape recorder.**

* * *

To judge whether your modified objective is acceptable, first see if it describes an observable student act or an observable product of a student act. If so, it is a behavioral objective. For example, the following few altered objectives would be acceptable.

**The student, when presented with the RCA tape recorder and a blank reel of tape, can record his own voice and play it back.**

**The student will list in writing the seven most important parts of the RCA tape recorder, and describe the use of each.**

**The student will point out six control switches on the RCA tape recorder and orally describe the way each is used.**

If your objective is similar to these, it can be considered a behavioral objective.

Here is a broader non-behavioral objective.

**The pupil will become familiar with the background of World War II.**

Change it on your answer sheet so that it is behavioral.

In the first place, you must now realize how many options are available in behaviorizing such an objective. The number of behavioral changes which might be made are almost unlimited. For instance, a teacher could have the students do anything from merely listing a few causes of World War II to more complex behavior such as writing a documented research paper. To tell whether your modified objective is properly stated, you will have to judge whether the objective describes an observable student behavior or an observable product of student behavior. If it does, your objective is acceptable.

You should now be able to distinguish between behaviorally and non-behaviorally stated objectives. You should also be able to change a non-behavioral objective to a behavioral objective. Why should these skills be of any benefit to you? In other words, what are the particular advantages of behavioral objectives?

For one thing, an instructor or product developer specifies his objectives in terms of student behavior. This enables him to select appropriate evaluation procedures, for there is less ambiguity with respect to the meaning of the objective.

In the same vein, behavioral objectives make it far easier for the teacher or programmer to select suitable learning activities for the class since he knows precisely what kind of student behavior he is attempting to produce.

Another advantage of behavioral objectives is that, since they are stated so specifically, the instructor himself can decide just how adequate his instructional objectives are. He can also call on a colleague to improve the quality of his objectives. When objectives are stated so broadly, it may be impossible to improve them. Specificity enables one to evaluate whether he is pursuing the right aims.

Here is another significant advantage of behavioral objectives. They can be given to mature students in advance of the instruction so that the student can focus his energies on relevant tasks. He can then avoid spending his time either mastering peripheral material or in trying to out-guess the instructor. It may take several sentences to state an objective clearly, but once this is done, both instructor and student can be guided properly.

Finally, behavioral objectives make it possible for the teacher, and others, to evaluate instruction on the basis of whether the students accomplish the intended objectives. No longer need an instructor be evaluated on whether he has a "pleasing personality" or a "wholesome philosophy of life". By employing behavioral objectives, teachers can chart their instructional goals and then go about accomplishing them. By gathering behavioral evidence,the teacher can tell whether he should retain or modify his instructional approach. If, for example, students fail to behave the way he wishes them to after instruction, he can change the learning activities. If students do acquire the behavioral changes he wishes, then perhaps he should expand or intensify the nature of his objectives.

Certainly there are other factors involved in good instruction, but precise behavioral objectives can markedly help most instructors to empirically augment the quality of their teaching. The key to improved instruction is indeed Behavioral Objectives.

INSTRUCTIONAL OBJECTIVES (ANSWER SHEET) W. James Popham

| | | | | |
|---|---|---|---|---|
| 1. | YES | | NO | |
| 2. | YES | | NO | |
| 3. | YES | | NO | |
| 4. | A | B | | |
| 5. | A | B | C | D |
| 6. | A | B | C | D |
| 7. | A | B | | |
| 8. | A | B | | |
| 9. | A | B | | |
| 10. | A | B | | |
| 11. | A | B | | |
| 12. | YES | | NO | |

13. _____

_____

_____

_____

## ANATOMY OF A BEHAVIORAL OBJECTIVE

The most useful form for stating objectives is to express them in terms of the behavior (both cognitive and observable) to be developed in the student and the content area (subject matter or concept) to which this behavior is to apply.

AUDIENCE

The **student** can . . . . . .

BEHAVIOR

Cognitive (analyze, recall, apply, etc.)
Observable (select, draw, match, write, etc.)

CONTENT

Subject Matter (Macbeth, lenses, NRA, equations, etc.)
Concepts (definitions, principles, theorems, etc.)

e.g. **The student will analyze an argument by commenting on the validity of the supporting statements with respect to the conclusion.**

Instructional Session Four: **Hierarchy of Learning Behaviors**

A. **OBJECTIVES**

The participant

1. can state the six levels of the Taxonomy of Educational Objectives and the major sub-categories of each.
2. can order given behaviors in increasing mental complexity.
3. can identify the taxonomic level of given behavioral objectives.
4. can write behavioral objectives requiring cognitive skills at all levels of the Taxonomy.

B. **PARTICIPANT'S PREPARATION:** Talmage's Instructional Objectives, Part III

C. **INSTRUCTIONAL PROCEDURES:** Discuss the ranking of selected verbs in order of their increasing mental complexity--use practice sheet on Classifying Cognitive Behaviors.

D. **PARTICIPANT'S PREPARATION:** Bloom's Taxonomy of Educational Objectives, (Chapter I)

E. **INSTRUCTIONAL PROCEDURES:**

1. Discussion of Chapter One of Bloom's Taxonomy.
2. Discussion of the six levels of the Taxonomy and examples of objectives of each.
3. Provide opportunities for participants to write objectives from their subject matter area at the different levels of the Taxonomy.
4. Group critique of behavioral objectives produced by procedure No. 3.

F. **MATERIALS:**

1. Practice Sheet in Classifying Behaviors
2. Bloom's Taxonomy of Educational Objectives
3. Talmage's Instructional Objectives, Part III

G. **EVALUATION:** Proficiency Quiz No. 4.

## TAXONOMY OF EDUCATIONAL OBJECTIVES[1]

Cognitive Domain

**1.00 KNOWLEDGE**

- 1.10 Knowledge of Specifics
  - 1.11 Knowledge of Terminology
  - 1.12 Knowledge of Specific Facts
- 1.20 Knowledge of Ways and Means of Dealing with Specifics
  - 1.21 Knowledge of Conventions
  - 1.22 Knowledge of Trends and Sequences
  - 1.23 Knowledge of Classifications and Categories
  - 1.24 Knowledge of Criteria
  - 1.25 Knowledge of Methodology
- 1.30 Knowledge of the Universals and Abstractions in a Field
  - 1.31 Knowledge of Principles and Generalizations
  - 1.32 Knowledge of Theories and Structures

**2.00 COMPREHENSION**

- 2.10 Translation
- 2.20 Interpretation
- 2.30 Extrapolation

**3.00 APPLICATION**

**4.00 ANALYSIS**

- 4.10 Analysis of Elements
- 4.20 Analysis of Relationships
- 4.30 Analysis of Organizational Principles

**5.00 SYNTHESIS**

- 5.10 Production of a Unique Communication
- 5.20 Production of a Plan, or Proposed Set of Operations
- 5.30 Derivation of a Set of Abstract Relations

**6.00 EVALUATION**

- 6.10 Judgments in Terms of Internal Evidence
- 6.20 Judgments in Terms of External Criteria

[1]Bloom, Benjamin S., et al, Taxonomy of Educational Objectives, David McKay Co., Inc., New York, 1968.

## SELECTED OBJECTIVES IN THE SOCIAL STUDIES – PRACTICE SHEET

**Directions:** Classify each of the following objectives according to its proper taxonomic level.

_____ 1. Student can identify terms related to finance.

_____ 2. Student can use his knowledge of Washington's political philosophy as expressed in his foreign policy speeches to analyze a series of statements and select those most representative of Washington's viewpoint.

_____ 3. Student will demonstrate his knowledge of Washington's Farewell Address by answering questions implied but not directly stated.

_____ 4. Student will demonstrate his knowledge of the term, "protective tariff," by applying it to current day situations.

_____ 5. Student can examine a table presenting data on the Industrial Revolution and translate this information into verbal statements.

_____ 6. Student can distinguish between "expansionist" and "anti-expansionist" statements contained in current periodicals.

_____ 7. Student can distinguish between moral and practical arguments put forth by the South to defend the Institution of Slavery.

_____ 8. Student will be able to analyze the various theories of the causes of American Revolution as propounded by various historians and evaluate their validity.

_____ 9. Student will determine the implications of the attack on Fort Sumter by comparing it with a given hypothetical situation.

_____ 10. Assuming that President Andrew Johnson had been removed from office, the student can predict the possible consequences of such an action to the federal government.

PRACTICE WORKSHEET

Listed below are pairs of cognitive behaviors. Select the behavior that involves skills of greater mental complexity.

1. a. memorize
   b. analyze

2. a. identify
   b. conclude

3. a. relate
   b. assess

4. a. infer
   b. recognize

5. a. conclude
   b. rephrase

6. a. recall
   b. distinguish

7. a. interpret
   b. synthesize

8. a. classify
   b. validate

9. a. design
   b. organize

10. a. estimate
    b. deduce

11. a. evaluate
    b. differentiate

12. a. select
    b. contrast

Instructional Session Five

## Purposes of Achievement Testing

A. OBJECTIVES

The participant

1. can state the relationship between objectives, learning activities and evaluation.
2. can state five purposes of achievement testing.
3. can discuss three benefits of pre and post testing.
4. can differentiate, in a given situation, whether the testing done is measurement or evaluation.
5. can discuss how proper evaluation procedures are crucial to individualized instruction programs.
6. will indicate, in discussion, that assignment of grades is a minor purpose of achievement testing.
7. will encourage other teachers in his district to utilize test results for purposes other than assigning a grade.

B. **INSTRUCTIONAL SESSION:** Review chart presented in Instructional Session One.

C. **PARTICIPANT'S PREPARATION:** Purposes of Achievement Testing.

D. **INSTRUCTIONAL SESSION:** Discuss the purposes of achievement testing.

E. **MATERIALS:** Purposes of Achievement Testing.

F. **EVALUATION:** Proficiency Quiz No. 5.

## PURPOSES OF ACHIEVEMENT TESTING:

1. Pre-testing helps you, the teacher, know where and how to begin.
2. Pre-testing frequently identifies those students in need of remedial help.
3. Pre-testing provides a basis for flexible grouping in particular skill areas.
4. Pre-testing coupled with post-testing tells you, the teacher, how much your students have improved.
5. Test information provides a basis for student self-assessment and motivation.
6. Test scores provide a basis for normative comparison with other students and with community, state, or even national norms.
7. Test information provides one means for communicating intelligently with parents about a child's progress in terms of the child's capacity.
8. Test data furnish a basis for detecting, and hence for attempting to remedy, certain weaknesses in the curriculum.
9. To determine the level of student achievement in the subject matter of the course at some point in time, in order to assign a grade.
10. To provide a continuous assessment of student learning throughout a definite time period, (such as one school year), in order to pinpoint the time when "changes" in learning start to take effect and when they peak. Also continuous assessment over a period of several school years may compare learning curves.

Instructional Session Six

## Techniques of Measurement

### A. OBJECTIVES

The participant

1. can list three different techniques of evaluating pupil progress.
2. can state advantages and limitations of the different techniques of evaluating pupil progress.
3. can indicate from a list of scoring procedures which are objective and which are subjective.
4. can name an evaluatory procedure appropriate for measuring a given type of behavior sought.
5. can demonstrate knowledge that all testing situations provide samples of student behavior rather than being completely comprehensive by selecting test items to measure skills within a given unit.
6. can construct and use a table of specifications in achieving balance in a test.
7. can analyze a standardized test by specifying the type of student behavior to be elicited by each item.
8. can, given a portion of a standardized test, construct a table of specifications for it.
9. will write test items or problematic situations to utilize the different methods of measurement to obtain valid samples of student behavior.

### B. PARTICIPANT'S PREPARATION:

Measuring the Attainment of Educational Objectives

### C. INSTRUCTIONAL SESSION:

1. Discuss objective and subjective scoring.
2. Discuss testing as providing a sample of a student's total behavior pattern.
3. Discuss the advantages and limitations of essay and objective tests.
4. Group critique of the appropriateness of test situations to measure certain types of behavior using Selecting the Appropriate Measurement Technique work sheet as a guide.
5. Group examination and discussion of the physics test and the accompanying Table of Specifications.

### D. PARTICIPANT'S PREPARATION:

Examine Iowa Test of Basic Skills and Teacher's Manual.

### E. INSTRUCTIONAL SESSION:

1. Delineation of the student behavior attempted to be elicited by each test item.
2. Provide practice in plotting student behaviors and items on Table of Specifications.
3. Discuss the usefulness of the Table of Specifications for obtaining balance in a test.
4. Discussion of the probable effectiveness of I.T.B.S. in eliciting these behaviors.

F. **MATERIALS:**

1. Measuring the Attainment of Educational Objectives
2. Iowa Test of Basic Skills and Teacher's Manual
3. Selecting the Most Appropriate Measurement Technique--Work Sheet
4. Table of Specifications
5. Physics test and the chart plotting it on a Table of Specifications
6. Assembling the Test

G. **EVALUATION:** Proficiency Quiz No. 6.

## MEASURING THE ATTAINMENT OF EDUCATIONAL OBJECTIVES

### I. Written Tests

#### A. ADVANTAGES AND LIMITATIONS

1. Characteristics of Essay Tests

   a. The student produces and organizes his own answers.
   b. The answer is in the student's words.
   c. The test is limited to a few items.
   d. Answers can be judged as to the degree of correctness.

2. Characteristics of Objective Tests

   a. The task is highly structured.
   b. The student selects from given alternatives.
   c. There is a large sample of items.
   d. One uses a predetermined key.

#### B. WRITING AND SCORING ESSAY ITEMS

1. Writing Essay Items

   a. Have in mind the mental process you want to be elicited by the student.
   b. Start with phrases like "compare," "contrast," "present arguments for and against," "give examples of."
   c. Define the task clearly.
   d. In achievement tests, do not use "in your opinion," etc.
   e. Beware of giving too many items in the time allowed.
   f. Do not mix essay and objective items when time is limited.
   g. Do not offer a choice of questions to be answered.

2. Scoring Essay Items

   a. Decide in advance what is to be measured.
   b. If more than one quality is being evaluated, give separate appraisals.
   c. Prepare a model answer with points to be covered and credit distribution.
   d. Read all answers to one question before going on.
   e. Grade as anonymously as possible.
   f. More than one rater improves reliability.

In subjective evaluation, the score assigned is determined by the personal opinion and judgment of the scorer. It is not determined by a prescribed scoring key or by the specifications in detail of the requirements for answers which will receive various scores. Objective evaluation utilizes a simple predetermined list of correct answers so that opinion or judgment in the scoring procedure is eliminated. The scoring of multiple choice, true-false and matching items is objective. Short answers or open-ended questions are partly objective, partly subjective and essay items are usually totally subjective.

#### C. WRITING AND SCORING OBJECTIVE ITEMS

1. General

   a. Keep the reading difficulty low (unless you are measuring reading or verbal skills).

b. If opinion or authority is cited, indicate WHOSE opinion or WHAT authority.
c. Be sure one item does not provide cues to another.
d. Avoid interdependent items.
e. Make the correct answer vary randomly among the choices in a set of items.
f. Avoid trick questions.
g. Avoid ambiguity.
h. Avoid trivia.

2. True or False

a. Be sure the item can be unequivocal and classified as true or false.
b. Avoid specific determiners.
c. Avoid indefinite terms of degree (most, some, few).
d. Avoid negative statements.
e. Keep the item to one idea.
f. Avoid items where the answer depends on one insignificant word, phrase or letter.
g. Avoid clueing by length of the item.

3. Short Answer

a. Avoid indefinite or open items.
b. Omit only key words.
c. Avoid a large number of blanks.
d. Blanks are better placed at the end than the beginning.
e. If a numerical answer is required, indicate the units in which it is to be expressed.

4. Matching

a. Items in a set should be homogeneous.
b. The number of answers should be greater than the number of problems.
c. The set should be relatively short.
d. Answer choices should be in a logical order, if one exists.
e. Directions should specify the basis for matching.
f. Directions should indicate if an answer can be used more than once.

5. Multiple Choice

a. The lead or stem should clearly formulate the problem.
b. Include as much as possible in the stem but avoid irrelevant material in the stem.
c. Be sure there is only one correct or best choice.
d. Items measuring application, etc. should be phrased in novel terms.
e. Avoid grammatical inconsistencies between the stem and alternatives.
f. Avoid "none of the above," "all of the above," "B & C," etc. unless they are absolutely necessary.
g. If the negative is necessary in the stem, capitalize or underline the "not" or "except", etc.

## II Oral Tests

Because it must be done individually, oral testing is demanding in terms of teacher time and classroom management. Nevertheless, oral testing is necessary where learners cannot read or have handicaps that do not allow them to write or mark answers.

### III. Performance Tests

This form of testing presents the student with a real situation and rates him on the procedure he uses to obtain a product or solve a problem, on the product or problem solution itself, or both. In some areas, such as physical education, vocal music and the dance, form is an aspect of performance important both to procedure and to product. Teachers of language, debate and speech may test performance in many situations. The ultimate test of learning in English Composition, for example, is the ability to write a passable composition in English. Instruction in French, Latin and Spanish aim at the spontaneous generation by students of acceptable sentences either orally or in writing.

**SELECTING THE MOST APPROPRIATE MEASUREMENT TECHNIQUE** Exercise Sheet

1. A chemistry teacher desires to measure his students' proficiency in the techniques of setting up a laboratory experiment. Describe the technique you believe would provide him with the most valid appraisal of their ability to do this task:

2. A social studies teacher is interested in measuring her pupils' progress in learning the skills connected with determining latitude and longitude. Describe the procedures that you believe would provide the most adequate index of their growth:

3. A language arts teacher is interested in measuring her pupils ability to compare and contrast the styles used by selected authors to achieve a certain reader response. What measurement techniques might be most appropriate to provide an index of the pupils' ability to do this task?

4. A social studies teacher desires to measure his pupils' mastery of selected historical concepts. His groups' skills to engage in reading and writing is very limited. What measurement technique might give the most valid estimate of their mastery of the concepts?

5. A science teacher wants to determine the level of his pupils' mastery of certain basic scientific data. Describe the techniques that he might use to obtain indications of their mastery of the basic data.

**TABLE OF SPECIFICATIONS**

Instructional
Objectives for:

| Grade level / Area of Application | | | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

**BEHAVIOR VS. CONTENT**

| | Knowledge of facts and Terminology | Understanding and Interpreting Graphical Material | Applying Principles to explain Phenomena | Recognizing Limitations of Principles | Evaluating Statements by Scientific Method |
|---|---|---|---|---|---|
| The Nature and Propagation of Light (16) | 1, 2, 3 5, 7, 13 14, 15 | | 21, 23 | 33 | 43, 44 45, 47 50 |
| Reflection and Refraction at a Plane Surface (9) | 17 | 41, 42 | 20, 24 30 | 31, 35 | 48 |
| Reflection and Refraction at a Single Surface (14) | 6, 8, 9 10, 11, 12 16 | 39, 40 | 22, 25 28 | 32 | 49 |
| Lenses (11) | 4, 18 | 36, 37 38 | 19, 26 27, 29 | 34 | 46 |
| | (18) | (7) | (12) | (5) | (8) |

PHYSICS TEST — OPTICS

1. The velocity of light in miles per second is about

   a. $3 \times 10^6$.
   *b. 186,000.
   c. 93,000,000.
   d. 8,000.

2. The astronomer who first discovered that light had a definite velocity was

   a. Galileo.
   b. Kepler.
   c. Newton.
   *d. Roemer.
   e. Hale.

3. Light waves resemble sound waves in that both

   a. are longitudinal.
   b. are transverse.
   c. can travel through a vacuum.
   *d. are forms of energy.

4. In a double concave lens, the image is

   a. erect, virtual and enlarged.
   b. inverted, real and diminished.
   c. erect, real and diminished.
   *d. erect, virtual and diminished.

5. Scientists now regard the principle of rectilinear propagation of light as

   a. entirely disproved.
   b. established without reservation.
   *c. valid for many practical purposes.
   d. valid in so few cases as to render it useless.
   e. of very little importance.

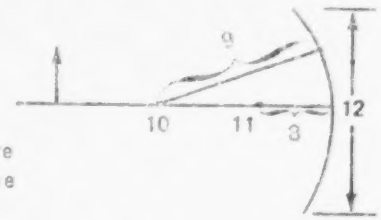
6. The focal length of a convex mirror

   a. is equal to R, the radius.
   *b. is always negative.
   c. is equal to 2R.
   d. is related to the aperture.
   e. is the point where all virtual images occur.

7. The property that light as a particle and light as a wave have in common is

   *a. reflection.
   b. diffraction.
   c. interference.
   d. infraction.

In the next five questions, identify the parts of the figure labeled by choosing the corresponding letter beside the correct name.



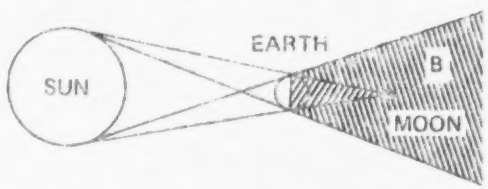

8. *b
9. *c
10. *d
11. *e
12. *a

a. aperture
b. focal length
c. radius of curvature
d. center of curvature
e. focal point

13. The ratio of the velocity of light in a vacuum to its velocity in any other medium is called the

a. index of reflection.
*b. index of refraction.
c. index of diffraction.
d. velocity ratio.
e. ratio of interference.

14. The section marked B in the figure below is the




a. corona
b. shade
*c. penumbra
d. reflection
e. umbra

15. Which part of the eye corresponds to the adjustable diaphragm of a camera?

a. pupil
b. lens
c. eyebrow
*d. iris
e. cornea

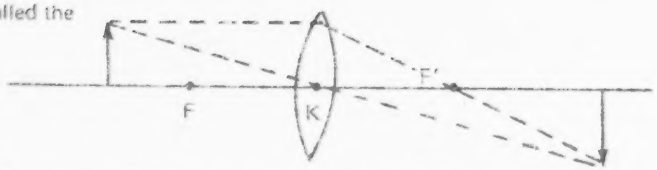
16. The failure of parallel rays to converge at the focal point of a spherical mirror is called

a. astigmatism.
b. declination.
c. chromatic dispersion.
*d. spherical aberration.

17. The angle between a ray of light striking a surface and the normal to that surface at the point is the

a. angle of coincidence.
*b. angle of incidence.
c. critical angle.
d. angle of total reflection.

18. In the figure below, K is called the




a. focal point.
b. lens centroid.
*c. optical center.
d. optical focus.

19. To determine quickly the focal length of a convex lens, one can focus light from a very _____ object upon a white screen.

*a. far
b. near
c. large
d. small
e. light

20. To hit a stick under water one should aim

a. directly at the stick.
b. in front of it.
*c. below it.
d. above it.
e. behind it.

21. An eclipse of the moon occurs when

*a. the earth is between the sun and the moon.
b. the moon is between the earth and the sun.
c. the sun and moon are at right angles with respect to the earth.

22. Given two examples of diffuse reflection, which is correct?

*a. A
b. B
c. A and B
d. neither A or B

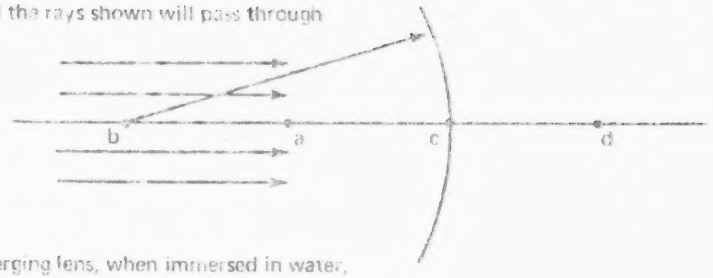
23. An airplane flying at a height of 100 ft. casts a shadow on the ground which is _____ a shadow of the same plane flying at 10,000 feet.

a. much smaller than
b. much larger than
*c. essentially the same size as

24. The angle between the incident ray and a plane mirror is $25^{o}$. The angle between the incident ray and the reflected ray in this case is

a. $65^{o}$.
b. $50^{o}$.
c. $115^{o}$.
*d. $130^{o}$.

25. Theoretically, all the rays shown will pass through




*a.
b.
c.
d.

26. A solid glass converging lens, when immersed in water,

a. has $f \gg f$ in air.
b. has $f = f$ in air.
*c. has $f \ll f$ in air.
d. becomes a diverging lens.
e. has no longer any focusing power.

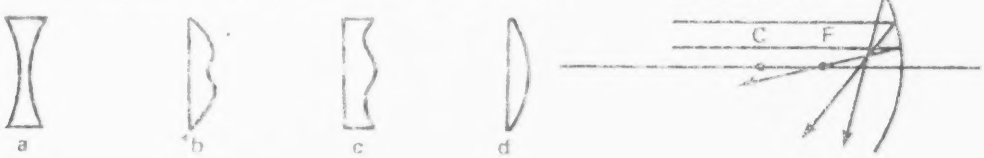
27. A convex lens formed by placing two watch glasses (each with parallel sides) edge to edge, taped together to make it airtight, if placed under water

a. will still be converging lens.
b. will not act like a lens at all.
*c. will act like a diverging lens.
d. will let rays pass through it unaffected.

28. A boy in front of a cylindrical mirror finds his image is

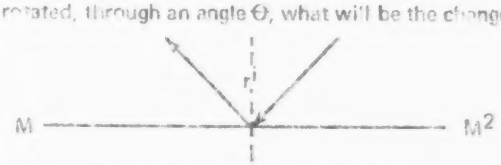
*a. extended in width.
b. diminished in width.
c. extended in height.
d. diminished in height.

29. Which lens would you use to get the rays to converge at F?




a
*b
c
d

30. If the mirror $MM^1$ is rotated, through an angle $\theta$, what will be the change in angle r?

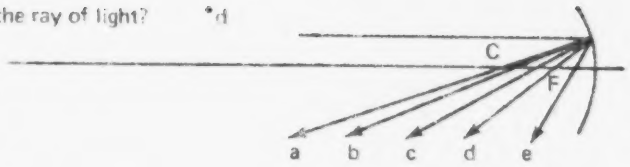



a. $\theta$
b. $\theta/2$
*c. $2\theta$
d. $4\theta$

31. Since a periscope depends on the relative refraction indexes of glass and air, does a periscope work if totally submerged? i.e., can you still see underwater with it?

a. No, because the water would bend the light rays out of range of the second prism.
b. Yes, because refraction in a periscope is unaffected by the two media.
c. No, because there is not enough light to cause sharp images.
*d. Yes, because the only refracting is done within the periscope where air is still the medium.

32. Which is the true path for the ray of light? *d




33. A pinhole camera does not work when

a. the object is too far away.
*b. the hole is made wider.
c. the hole is not centered in the camera.
d. the screen is black.

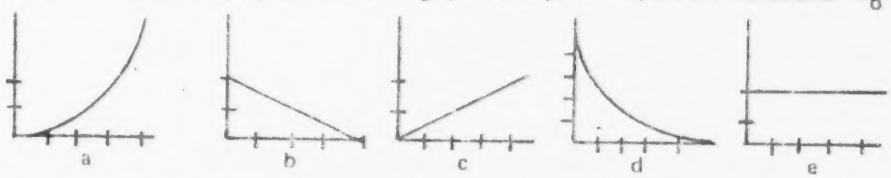
34. If two lenses have identical shapes and sizes,

a. f is independent of n.
b. $f_1 = f_2$.
c. $n_1 < n_2$ implies $f_1 < f_2$.
*d. $n_1 > n_2$ implies $f_1 < f_2$.

35. Two boys standing on a beach with a straight shoreline, try to make the beam of a flashlight held by one reflect off the still water back to the other.

a. It is possible to accomplish this because i = r.
b. It is possible because light travels in straight lines.
c. It is impossible because $i \neq r$.
*d. It is impossible because the incident ray, the reflected ray and the normal must be in the same plane.

Each question is answered by one of the five graphs below, For a - e, the horizontal axis = $S_o$




An object is placed at the principle focus of a converging lens and moved slowly from the lens.

36. Which graph could represent the location of the image as the object moves away? *d

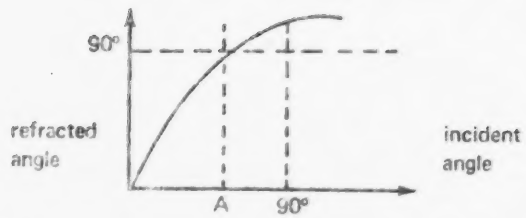
37. Which graph indicates the product of $S_o$ and $S_i$ as the object moves away? *e

38. Which graph could represent the height of the image as the object moves away? *d

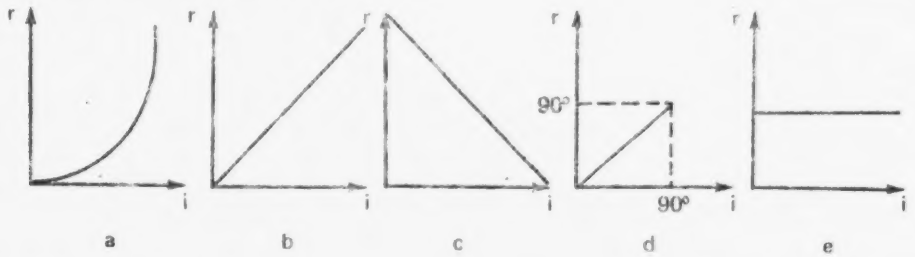
39 The object is now placed at the principle focus of a parabolic mirror and moved away from the mirror; which graph could represent the ratio $h_o / h_i$ as the object moves away? *c

40. Suppose an object moves along the principle axis toward the principle focus of a parabolic mirror at a speed of 1 m/sec. Which graph would represent the speed with which the image would move away from the principle focus as a function of time? *a

41. In the figure, angle A is called
    a. the reflected angle.
    b. the normal angle.
    *c. the critical angle.
    d. the acute angle.




42. Which graph is correct for reflection? *d




For Questions 43 - 50 express your evaluation of the statement by choosing the corresponding number of the correct stage.

a. Belief Stage
b. Authoritative Opinion Stage
c. Observation Stage
d. Controlled Experiment Stage

43. The moon reflects light from the sun. *c

44. White light is a combination of the different colors of the sun's spectrum. *d

45. Light travels from place to place instantaneously *a

46. Convex lenses magnify small objects. *d

47. The speed of light is the fastest thing obtainable in the universe. *b

48. The angle of incidence equals the angle of reflection. *d

49. A ray of light is bent more by glass than water. *d

50. Light is a particle. *c

## ASSEMBLING THE TEST

**LENGTH –** The test should be long enough to sample significant behaviors called for by each objective, but short enough so that the student can complete it comfortably within the time allotted.

**ITEM ORDER –** One useful way to order test items is in accordance with the sequence of the content itself. Also, items may be arranged in the order of their perceived difficulty.

**KEY –** Prepare the answer key before the test is duplicated so that the correct responses can be randomly keyed.

**DIRECTIONS –** Keep directions simple and explicit.

**LANGUAGE –** Language should be direct and precise. The student should be able to answer on the basis of his information and understanding of the subject matter alone.

**APPEARANCE –** Avoid overcrowded, or poorly organized format. Be sure the test is proofread for errors before the student gets it and that the errors are corrected.

**ITEM TYPE –** The type of item to be used should be carefully selected so that it is the best possible way of measuring the degree of pupil attainment of each instructional objective. No more than two or three different kinds of items should be used in one test and they should be grouped together.

The above list by no means exhausts the important factors in putting together a good objective test. You will probably think of others. These suggestions were taken from Green, John A., Teacher-Made Tests, Harper and Rowe, New York, 1963 and Gorow, Frank F., Better Classroom Testing, Chandler Publishing, San Francisco, 1966.

Instructional Session Seven:

## Multiple - Choice Items

A. OBJECTIVES

The participant

1. can list at least three advantages and two limitations of multiple-choice items.
2. can, given a population of students, state the appropriate number of alternatives for items to be presented.
3. can state three advantages of objective scoring as it relates to multiple-choice items.
4. can state the definition of terms used in discussing multiple-choice items, e.g., stem, foil, etc.
5. can, given a multiple-choice item, identify the different parts.
6. can, given the different forms of multiple-choice items, name each.
7. can identify faults in given multiple-choice items.
8. can rewrite correctly "wrongly stated" multiple-choice items.
9. can write multiple-choice items that are free of item faults.
10. can write multiple-choice items that measure cognitive processes more complex than simple recall of specific facts.
11. can write multiple-choice items that require the student to perform a task critical to the performance of the objective it was written to test.
12. will listen openmindedly to the advantages and methods of writing multiple-choice items as evidenced by participation in discussion.
13. will provide evidence of believing that multiple-choice items can be used to test higher mental processes than recall of specific facts by actually attempting to write such items.
14. will express to others the advantages and scope of objective testing, especially with multiple-choice items.
15. will encourage others to use multiple-choice items in evaluating the degree of attainment of behaviorally stated objectives.

B. PARTICIPANT'S PREPARATION: Advantages and Limitations of Multiple-Choice Items.

C. INSTRUCTIONAL SESSION:

1. Discuss the Advantages and Limitations of Multiple-Choice Items.
2. Distinguish clearly between product and process.

D. PARTICIPANT'S PREPARATION: Anatomy of the Multiple-Choice Item.

E. INSTRUCTIONAL SESSION: Review briefly the material presented in The Anatomy of the Multiple-Choice Item.

F. PARTICIPANT'S PREPARATION:

1. Typical Item Faults

2. Examples of Item Faults – Practice Sheet

G. **INSTRUCTIONAL SESSION:**

1. Review participant's responses on the Example of Item Faults practice sheet.

2. Group critique participants' efforts to correct item "faults" on Practice Sheet.

3. Provide practice in writing multiple-choice items for the behavioral objectives written by the participants in Instructional Session Three, B–6.

4. Group critique items produced by participants through activity No. 3.

H. **MATERIALS:**

1. Advantages and Limitations of Multiple-Choice Items.

2. Anatomy of the Multiple-Choice Item

3. Typical Item Faults

4. Examples of Item Faults – Practice Sheet

I. **EVALUATION:** Proficiency Quiz No. 7.

## ADVANTAGES AND LIMITATIONS OF MULTIPLE–CHOICE ITEMS

Well constructed multiple-choice items have certain advantages for the student and the teacher as a means of measuring the attainment of instructional objectives.

A. ADVANTAGES OF MULTIPLE–CHOICE ITEMS:

1. The student must demonstrate the specific ability the item calls for in order to select the correct response. He cannot avoid direct confrontation with the question.
2. Multiple-choice items may be objectively scored since such factors as skill in verbal expression, or speed and legibility of handwriting do not affect the teacher's assessment of the student's response. Only the response to the question enters into the score.
3. Since multiple-choice items are relatively brief, they allow a wide sampling of behaviors in a short time. Writing (composing and recording) answers is generally conceded to be more time-consuming than reading and selecting an answer.
4. Each multiple-choice item has one, stated, correct answer. Thus, scoring can be done by almost anyone, using a prepared key.
5. Scoring of multiple-choice items is fast and easy because the student's response is simple and unambiguous. Multiple-choice items may be machine scored.
6. Multiple-choice items can be analyzed statistically with much greater ease than free response items.
7. Many kinds and levels of learning behavior may be tested by multiple-choice items.
8. The multiple-choice item generally provides greater test reliability per item than the true-false item.

B. LIMITATIONS OF MULTIPLE–CHOICE ITEM:

1. Handwriting skills, performances on musical instruments and similar skills and abilities should not be tested with multiple-choice kinds of tests.
2. Although multiple-choice items can efficaciously measure the higher mental abilities, considerable caution must be exercised to insure that the cause of the pupil's lack of ability to engage in the higher mental processes as measured by the individual test item is not just the forgetting of certain facts. This limitation can be overcome by two methods:
   a. Grouping items according to varying mental complexities on a specific topic so a more adequate diagnosis can be made of the pupil's real problem.
   b. Including the necessary information in the test situation to insure that the pupil's lack of ability to engage in the higher mental processes is not strictly the lack of knowledge of selected facts.

Ebel, Robert L., Measuring Educational Achievement, Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1965
Furst, Edward J., Constructing Evaluation Instruments, New York, David McKay Company, Inc., 1958.

## THE ANATOMY OF THE MULTIPLE–CHOICE ITEM

The multiple-choice item is an example of a supplied-response test. Each multiple-choice item has two (2) parts. These are:

### A. STEM OR LEAD

1. Directions are a part of the stem or lead. Sometimes the lead is preceded by directions telling what to do in order to respond correctly.

2. The stem or lead is of two main types.

   a. The direct (or bona fide) question
      Example - Why do living organisms need oxygen?

   b. The incomplete statement
      Example - During swallowing, the epiglottis covers the opening of the larynx so that _____.
      Note: This type of lead is more difficult to write correctly but if it is well formed, it usually results in a more concise and pointed lead.

### B. ALTERNATIVES

These alternatives are often called choices, options, or answers.

1. One answer is clearly best and is the correct response.

2. The "crippled" or inferior alternatives are called distractors (or decoys or foils) and are incorrect answers.

3. The alternatives may be expressed as words, clauses, sentences, numbers, pictures, graphs, diagrams and formulae; in short, anything that satisfies the logical requirements of the question posed is allowable.

4. There are usually 4 or 5 alternatives. Items appropriate for the primary and intermediate grades may offer only 3 alternatives.

**TYPICAL ITEM FAULTS – Revised Material**

1. Pupil behavior attempted to be elicited by the item is **not** congruent with behavior stated in the objective.

   For Example: Objective: The student can recognize sketches, drawings, or pictures of the past five presidents of the U. S.

   Test Item: Which of the four men listed below *was* never President of the U. S.?

   a. George Washington
   b. Robert Kennedy
   c. Lyndon Johnson
   d. Harry S. Truman

2. Item's stem does **not** present a **single, definite, important** question or problem.

3. Item's alternatives do **not** include a **single, correct,** or **best** answer or completion.

4. If the form of the stem is an incomplete statement, it should contain as much of the problem as possible.

   For example: Washington

   a. was a great general.
   b. was a plantation owner.
   c. was the first President of the U. S.

   Improved: George Washington is often referred to as the "Father of our Country" because he was

   a. a great general.
   b. a plantation owner
   c. the first President of the U. S.

5. Item's alternatives are **not** stated as briefly as possible.

6. Question or problem is **not** stated in positive terms. (If it is necessary to use a negative, it should be underlined and/or capitalized, e.g. **NOT, EXCEPT, UN**likely.)

7. Alternatives are **not** grammatically consistent with the lead.

8. Alternatives are not sufficiently similar in grammatical construction, degree of precision, or length.

9. Alternatives are **not** plausible. Incorrect responses should tempt pupils who have incomplete or superficial knowledge of the material.

10. Alternatives contain a "trick" word in an otherwise correct response just for the purpose of catching the reader.

## EXAMPLES OF ITEM FAULTS – PRACTICE SHEET (Revised Materials)

Each of the following items contain one or more of the faults listed on the sheet, Typical Item Faults. You are to do two things with each of the items:

a. List the number(s) of the fault(s).

b. Correct the item so that it is acceptable.

1. Which of the following has helped most to increase the average length of human life?

   a. fast driving

   b. avoidance of over-eating

   c. wider use of vitamins

   d. wider use of inoculations

2. At the Constitutional Convention, the "great" compromise

   a. gave small and large states equal representation in the Senate.

   b. made slave holding illegal.

   c. was opposed by Washington.

   d. gave western lands claimed by the states to the federal government.

3. The part of speech that tells "how", "when", or "where" is an

   a. noun

   b. pronoun

   c. adverb

   d. conjunction

   e. verb

4. The framers of the Constitution faced many problems. The delegates to the Constitutional Convention represented states with different interests and the delegates from the individual states wanted to see that their states' interests were protected. However, the delegates agreed that the Articles of Confederation needed to be changed in order to provide for

   a. a president whom everyone could respect.

   b. a stronger central government.

   c. a better understanding between states.

   d. a government for and by the people.

5. Which one of the following leaders of the Revolutionary War did not want the Articles of Confederation changed?

   a. Benjamin Franklin

   b. George Washington

   c. Alexander Hamilton

   d. Patrick Henry

6. The Chicago Symphony

   a. presents the largest number of concerts.
   b. has the largest number of players.
   c. has the largest number of male players.
   d. has the largest number of female players.

7. Slavery was first started

   a. at Jamestown settlement.
   b. at Plymouth settlement.
   c. at the settlement of Rhode Island
   d. a decade before the Civil War

8. Which event in the following list has been of the greatest importance in American history?

   a. Braddock's defeat
   b. Burr's conspiracy
   c. The Hayes-Tilden contest
   d. The Webster-Hayne debate

9. The adoption of the Constitution was generally opposed by people who

   a. owed money
   b. thought that most people were unfit to govern themselves
   c. owned businesses
   d. thought the states should be destroyed

10. The student knows that (1) the Constitution provides for a census every ten years and (2) the reason that a provision was made for a regular national census.

    A census every ten years was provided for in the Constitution in Article I, Section

    a. 1
    b. 2
    c. 3
    d. 4

Instructional Session Eight:

## Basic Statistical Measures

A. OBJECTIVES

The participant

1. can demonstrate knowledge of the definitions of certain statistical terms (mean, median, mode, skewness, kurtosis, normal distribution, standard deviation, and correlation) by matching correct pairs from lists of terms and definitions.
2. can calculate for a given distribution of a small sample, the mean, mode, median, variance and standard deviation.
3. can determine whether a given distribution is negatively or positively kurtotic.
4. can determine whether a given distribution is negatively or positively skewed.
5. can, by giving examples, demonstrate a knowledge of the implications of positive, negative, and zero correlation.
6. can, by giving an illustrative example, show that correlation does not necessarily imply causality.

B. PARTICIPANT'S PREPARATION: Basic Statistical Measures - Numerical Example (2 pages)

C. INSTRUCTIONAL SESSION:

1. Discuss definition of frequency distribution and histogram.
2. Discuss problems in selecting intervals and the normal curve.
3. Discuss definitions and formulae for mode, median, mean, variance and standard deviation.
4. Calculate the actual values for the given data for each term in 3.

D. PARTICIPANT'S PREPARATION: Basic Statistical Measures - Samples of Skewness and Kurtosis

E. INSTRUCTIONAL SESSION

1. Discuss positive and negative skewness (definition and example)
2. Discuss positive and negative kurtosis (definition and example)
3. Discuss meaning of significant skewness or kurtosis relative to normal curve

F. PARTICIPANT'S PREPARATION: Standard Score Scales

G. INSTRUCTIONAL SESSION

1. Discuss role of standard scores, in terms of standard deviations and the advantages over raw score reporting.
2. Discuss the relation of some of the familiar scales to standard scores.

H. PARTICIPANT'S PREPARATION:

1. Basic Statistical Measures - Examples of Correlation

2. Scores For a Fifth Grade Class (2 pages)

I. **INSTRUCTIONAL SESSION**

1. Discuss correlation as a measure of relation between two variables; meaning of positive and negative correlation.
2. Discuss the notion that correlation does **not** imply causality.
3. Using the scores and the formula for correlation, discuss the calculations that produce the correlations and whether they conform to intuitive feelings about the variables.

**MATERIALS:**

1. Basic Statistical Measures - Numerical Examples
2. Basic Statistical Measures - Samples of Skewness and Kurtosis
3. Basic Statistical Measures - Examples of Correlation
4. Standard Score Scales
5. Scores for a Fifth Grade Class
6. Glossary of Measurement Terms

J. **PROFICIENCY QUIZ NO. 8**

## FREQUENCY DISTRIBUTION

The frequency distribution of a set of data (e.g., test scores) can be represented in a variety of ways. Two are pictured in the NUMERICAL EXAMPLE. The first is merely to state the number of scores that occur in a set of intervals of scores where the intervals are determined by the number and range of scores. The second way to represent a frequency distribution is by means of a histogram or bar graph. The vertical axis represents the number of scores, while the horizontal axis is the score itself or intervals of scores.

## THE NORMAL CURVE

When the frequency distribution has a bell-like shape centered about the mean, it is said to be "normal". (It is unfortunate that the word "normal" carries connotations that are irrelevant to statistics.)

Normal distributions are very common. One has only to consider amounts like the number of leaves on oak trees versus the number of oak trees to conclude that many trees have nearly the same number of leaves as the average number. Few have *few* leaves or an enormously large number. Intelligence is another example with great numbers of people near 100 and few people at the extreme ends of the scale.

No classroom teacher should be satisfied with her accomplishments if her students' test scores are distributed normally. It is much more desirable to have a large proportion of students at the upper end of the scale which would result in a shape far from bell-like.

## STANDARD DEVIATION

Knowing only the mean or average of a set of scores does not allow us to make any judgments about differences among the subjects tested; i.e., is the group fairly homogeneous or not? Some measure of the spread of scores is necessary to answer such a question. One measure of spread is the standard deviation.

A seemingly naive approach to measuring the spread of test scores would be to take the average distance of each test score from the mean. If the scores differ greatly, this value would be large. If the scores were quite similar, the result would be small.

For example, consider the following scores for nine students on a ten-item test.

1, 3, 4, 4, 6, 6, 6, 7, 8

The average of these scores is 5. Now let us calculate the distance from each score to the mean.

$1 - 5 = -4$
$3 - 5 = -2$
$4 - 5 = -1$
$4 - 5 = -1$
$6 - 5 = 1$
$6 - 5 = 1$
$6 - 5 = 1$
$7 - 5 = 2$
$8 - 5 = 3$

Now we shall take the sum of these and find the average "distance" from the mean.

Sum = 0 $\qquad$ 0/9 = 0

Summing these distances produced a result of zero. Actually, this is not too surprising since the mean is defined as the "center of gravity" for all scores considered.

Statisticians avoid this problem by averaging the **squares** of the distances from the mean and later taking the square root of the average calculated as the measure of spread.

Therefore

$$(1-5)^2 = -4^2 = 16$$
$$(3-5)^2 = -2^2 = 4$$
$$(4-5)^2 = -1^2 = 1$$
$$(4-5)^2 = -1^2 = 1$$
$$(6-5)^2 = 1^2 = 1$$
$$(6-5)^2 = 1^2 = 1$$
$$(6-5)^2 = 1^2 = 1$$
$$(7-5)^2 = 2^2 = 4$$
$$(8-5)^2 = 3^2 = 9$$

The sum of these squared distances is 38, and the average equals 38/9 or 4.22. The standard deviation is the square root of 4.22 or 2.054.

With a class of 35 students, this method of squaring distances gets quite out of hand and another formula is used which arrives at the same number.

$$\text{standard deviation} = \sqrt{\frac{\Sigma x^2 - (N \times M^2)}{N}}$$

where $\Sigma x^2$ is the sum of each score squared

N is the number of scores

and $M^2$ is the mean squared.

This formula is applied to the numerical example in BASIC STATISTICAL MEASURES – Calculation of Mean, Mode, Median, Variance and Standard Deviation.

## STANDARD SCORE SCALES

When a group of data is normally distributed, scores can be thought of in terms of amount of standard deviations from the mean.

For example, suppose in a certain set of scores,

Mean = 75

St. Dev. = 10

then a score of 85 would be **one** standard deviation above the mean or a "standard score" of +1. A score of 60 would be **one and a half** standard deviations **below** the mean or a standard score of −1.5. (Standard scores are also called "Z scores.")

In a normally distributed set of scores, a **standard score** yields much more information than a raw score. This is evidenced by consulting the page entitled, STANDARD SCORE SCALES. Note that 68.26%

(34.13 + 34.13) of a group lies between one standard deviation below the mean and one standard deviation above the mean. Or, a score equivalent to one standard deviation above the mean corresponds to the 84th percentile. A score equal to the mean corresponds to a standard score of 0 or the 50th percentile.

The charts below the normal curve show the correspondence between score of the more commonly used tests; e.g., College Entrance Examination Board, Stanines, Percentiles, etc.

## SKEWNESS AND KURTOSIS

When a distribution of test scores is not bell-shaped or "normal," the terms used to describe the type of distortion from normality are **skewness** and **kurtosis.**

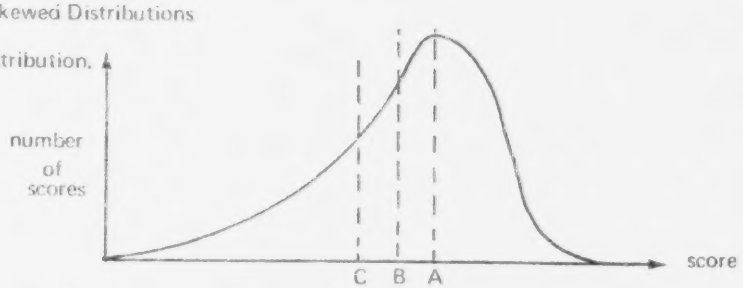
Skewness represents the imbalance to the right or left. When a large proportion of subjects is scored near the top of the distribution, the curve still looks bell-like near the upper (or positive) end. But at the left (or negative) end of the curve, instead of the bell shape, only a gentle sloping curve is seen. Thus, when a **large proportion receives high scores,** the distribution is **negatively skewed.** In a similar manner, when a large proportion receives **low scores,** the distribution is **positively** skewed. It is sometimes convenient to recall that skewness means distortion and the curves are distorted where the least number of scores occur.

Kurtosis refers to the homogeneity of the group represented in the distribution curve as a high, thin hill or a low flat hill. When the curve is high and thin (many subjects in a narrow score range) the kurtosis is positive. When the curve is low and flat, the kurtosis is negative; i.e., equal numbers of subjects over a wide range of scores.

For examples of positively and negatively skewed as well as positively and negatively kurtotic curves, see SAMPLES OF SKEWNESS AND KURTOSIS.

Mean, Mode and Median in Skewed Distributions

Consider the following distribution.




The mode is easily found, for this is the score which the greatest number of people received. (A). Here most of the distribution is in the higher scores so generally the median is higher than the mean. In addition, one or two very low scores greatly distort the mean. In the sketch above, C corresponds to the mean and B the median.

## CORRELATION

Correlation is a number representing the relation between two variables. Two tests are correlated if a score on the second can be predicted, knowing the score on the first. For example, word knowledge and reading achievement would be **correlated positively,** for if word knowledge was high, in all probability, reading achievement would also be high, and if one was low, the other would be expected to be low. Two variables are **correlated negatively** if they vary in opposite directions to each other. One example might

be discipline points and grade average. The more discipline points, the lower we would expect to find the grade average and the higher the grade average, the less we expect to see discipline points. Variables are **uncorrelated** if there is no relation between them; e.g., shoe size and score on a physics test. Here, knowing the value of one does not help in predicting a value on the other.

Correlation has the limits −1.0 and +1.0. The limit −1.0 corresponds to a "perfect negative correlation" or an opposite direction relation, while +1.0 corresponds to a "perfect positive correlation" or a same direction relation.

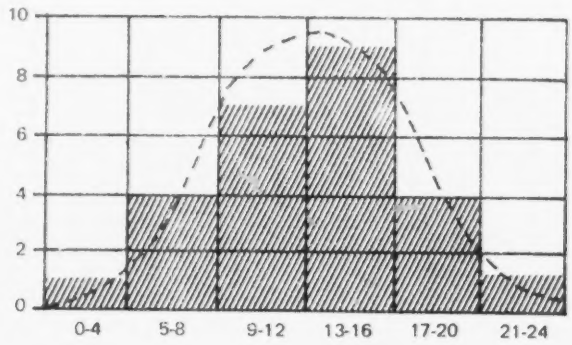
Some examples of correlations are given in EXAMPLES OF CORRELATION. Each person is represented by an "X" and graphed with respect to the two variables used.

Calculation of the correlation between two variables is illustrated in SCORES FOR A FIFTH GRADE CLASS.

Raw Data (Test Scores) for 26 students

4, 5, 6, 7, 8, 9, 9, 10, 10, 10, 11, 12, 13, 13, 14, 15, 15, 15, 15, 16, 16, 17, 17, 19, 20, 23

| Score Interval | Frequency |
|---|---|
| 0 – 4 | 1 |
| 5 – 8 | 4 |
| 9 – 12 | 7 |
| 13 – 16 | 9 |
| 17 – 20 | 4 |
| 21 – 24 | 1 |
| | 26 |



## BASIC STATISTICAL MEASURES

Calculation of mean, mode, median and variance, standard deviation

| $x$ | $x^2$ |
|---|---|
| 4 | 16 |
| 5 | 25 |
| 6 | 36 |
| 7 | 49 |
| 8 | 64 |
| 9 | 81 |
| 9 | 81 |
| 10 | 100 |
| 10 | 100 |
| 10 | 100 |
| 11 | 121 |
| 12 | 144 |
| 13 | 169 |
| 13 | 169 |
| 14 | 196 |
| 15 | 225 |
| 15 | 225 |
| 15 | 225 |
| 15 | 225 |
| 16 | 256 |
| 16 | 256 |
| 17 | 289 |
| 17 | 289 |
| 19 | 381 |
| 20 | 400 |
| 23 | 579 |

Mode = most frequent score = 15

Median = the score that divides the population into two equal parts = 13

$$\text{Mean} = \frac{\text{sum of scores}}{\text{no. in group}} = \frac{329}{26} = 12.65$$

$$\text{Variance} = \frac{\text{sum of squared scores} - [(\text{no. of scores}) \times (\text{mean}^2)]}{\text{number of scores}}$$

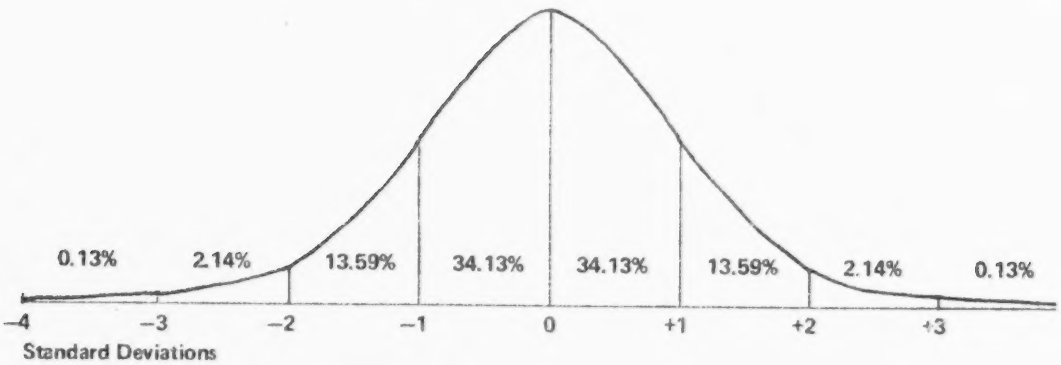
$$= \frac{\Sigma x^2 - N(M^2)}{N}$$

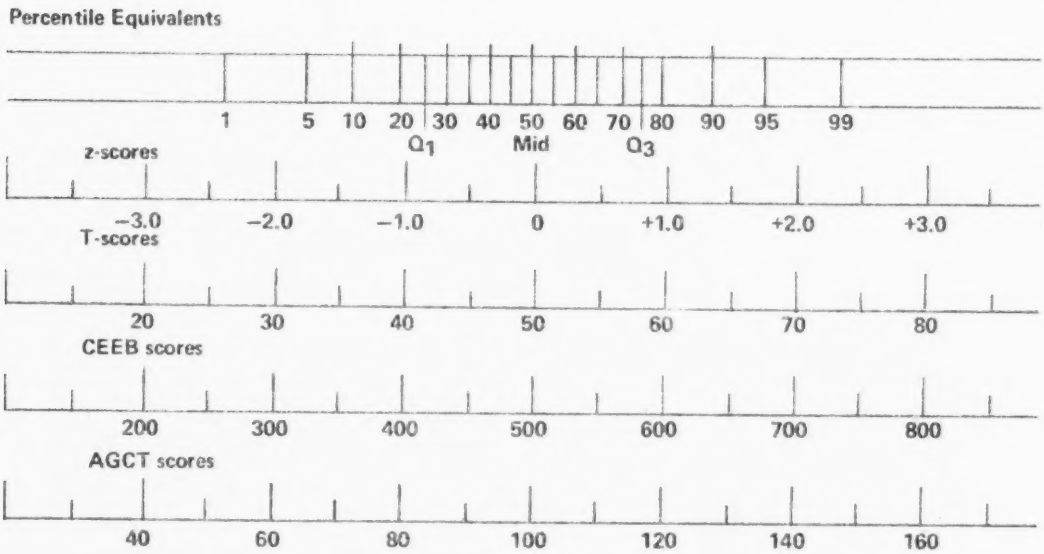
$$= \frac{4731 - 26(12.65)^2}{26}$$

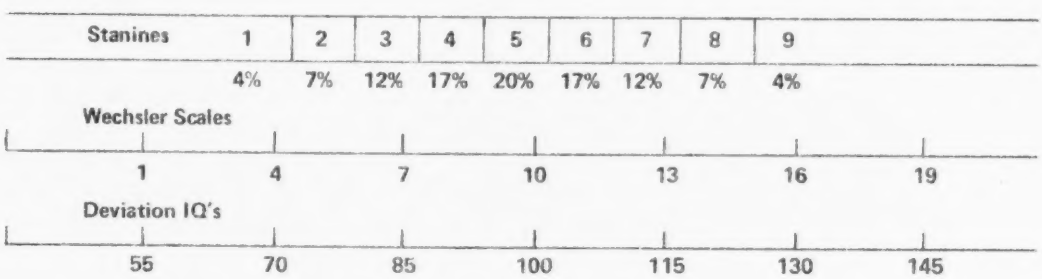
$$= 21.96$$

$$\text{Standard deviation} = \sqrt{\text{Variance}} = 4.69$$

## STANDARD-SCORE SCALES

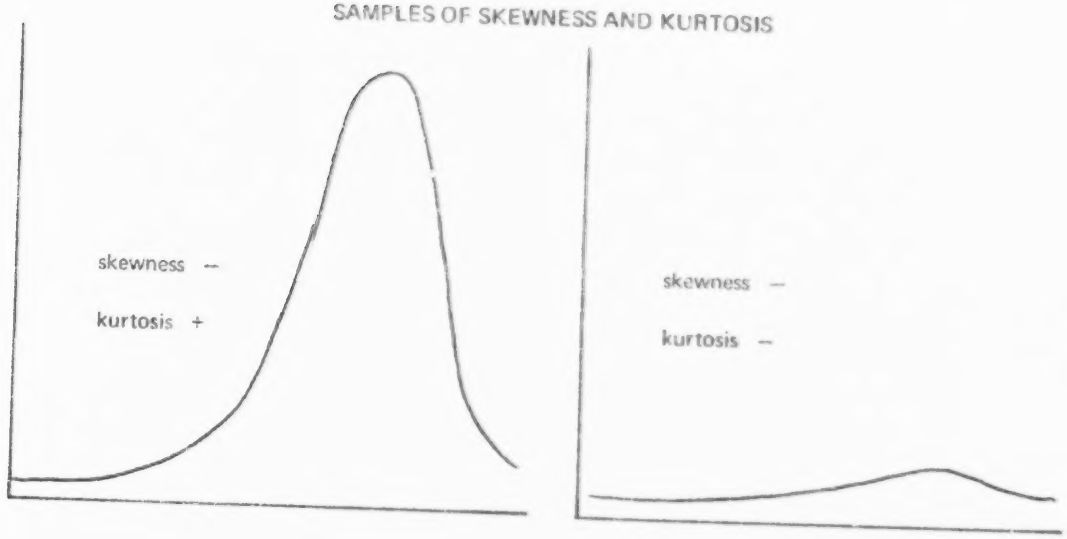


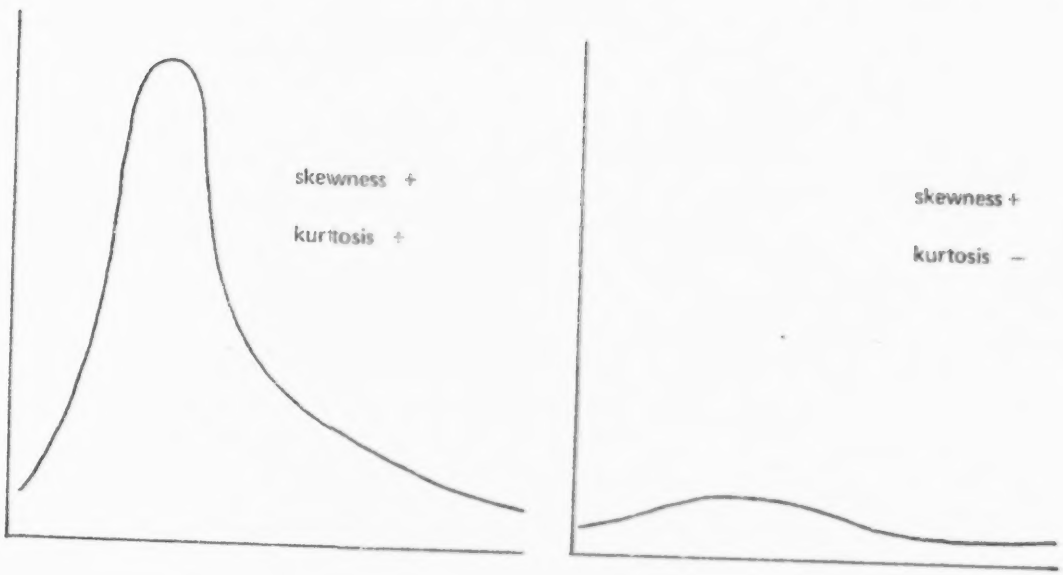



## BASIC STATISTICAL MEASURES

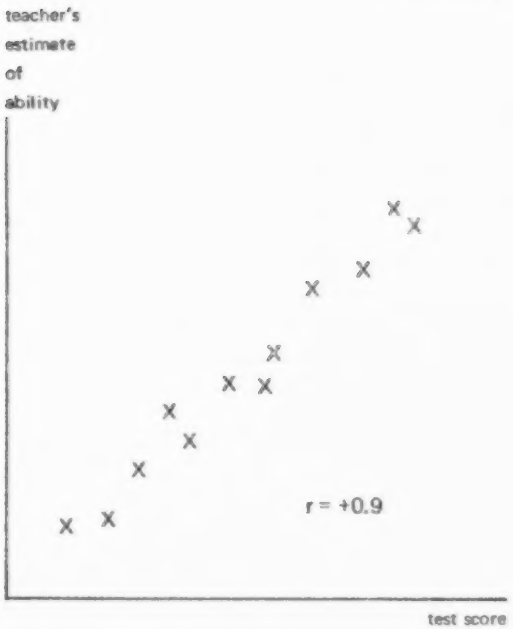
### SAMPLES OF SKEWNESS AND KURTOSIS

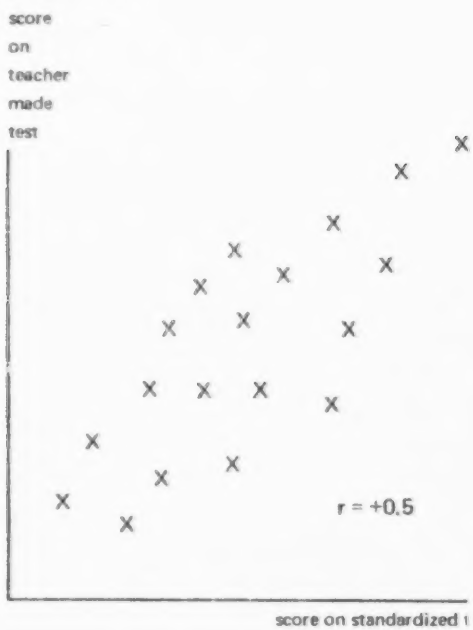


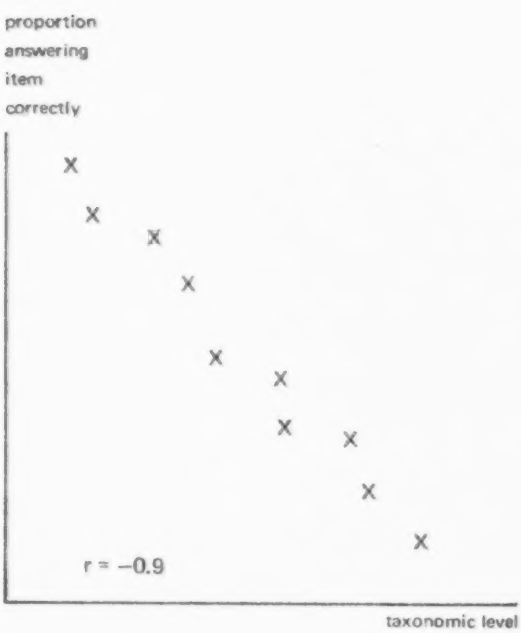



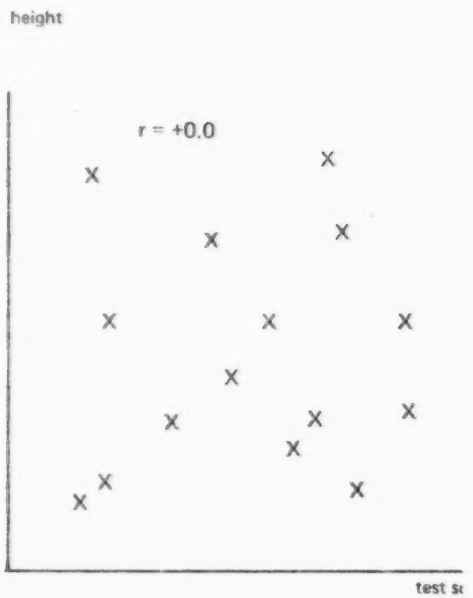
## BASIC STATISTICAL MEASURES

### EXAMPLES OF CORRELATION



SCORES FOR A FIFTH GRADE CLASS

| Test Score | Teacher's estimate of ability | Test standardized score | Discipline points | Height (inches) |
|---|---|---|---|---|
| 2 | D (1) | 3.2 | 10 | 60 |
| 5 | C (3) | 3.5 | 7 | 50 |
| 8 | B (4) | 4.1 | 8 | 52 |
| 10 | B (4) | 4.6 | 9 | 62 |
| 14 | B (4) | 5.0 | 6 | 49 |
| 15 | C (3) | 6.0 | 5 | 59 |
| 15 | A (5) | 5.8 | 3 | 63 |
| 16 | C (3) | 6.1 | 0 | 51 |
| 20 | C (3) | 6.0 | 2 | 53 |
| 25 | B (4) | 7.2 | 1 | 58 |

$$r_{xy} = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2]\,[N\Sigma y^2 - (\Sigma y)^2]}}$$

| Score | Teacher's Estimate | Standardized | Discipline Points | Height |
|---|---|---|---|---|
| $\Sigma x = 130$ | $\Sigma y = 34$ | $\Sigma z = 51.5$ | $\Sigma w = 51$ | $\Sigma s = 557$ |
| $\Sigma x^2 = 2120$ | $\Sigma y^2 = 126$ | $\Sigma z^2 = 280.15$ | $\Sigma w^2 = 369$ | $\Sigma s^2 = 31273$ |
| | $\Sigma xy = 473$ | $\Sigma xz = 747.30$ | $\Sigma xw = 478$ | $\Sigma xs = 7248$ |

$$\frac{10(473) - (130)(34)}{\sqrt{[10(2120) - 130^2][10(126) - 34^2]}} = \frac{310}{\sqrt{(4300)(104)}} = \frac{310}{668.7} = +0.46$$

$$\frac{10(747.3) - (130)(51.5)}{\sqrt{[10(2120) - 130^2][10(280.15) - 51.5^2]}} = \frac{778}{\sqrt{(4300)(149.25)}} = \frac{778}{801.1} = +0.97$$

$$\frac{10(478) - (130)(51)}{\sqrt{[10(2120) - 130^2][10(369) - 51^2]}} = \frac{-1850}{\sqrt{(4300)(1089)}} = \frac{-1850}{2164} = -0.85$$

$$\frac{10(7248) - (130)(557)}{\sqrt{[10(2120) - 130^2][10(31273) - 557^2]}} = \frac{70}{\sqrt{(4300)(2481)}} = \frac{70}{3266} = +0.02$$

| correlation of test score with | teacher's estimate of ability | standardized test score | discipline points | height (in inches) |
|---|---|---|---|---|
| | .46 | .97 | −.85 | .02 |

Instructional Session Nine

## Test Analysis Interpretation of Computer Output

A. **OBJECTIVE**

The Participant

1. can match with their definitions in given lists: item difficulty, item discriminating power, corrected score.
2. can determine a unique six character or less identification code for the subjects to be tested.
3. can find the most difficult and easiest items on a test using the item analysis results.
4. can suggest reasons for an item's extreme easiness or difficulty.
5. can cite examples where distractors were not selected and pose possible reasons for this.
6. can suggest reasons why a large proportion of the subjects would choose the same wrong alternative.
7. can give arguments for keeping or omitting an extremely easy or difficult item.
8. can apply the "rule of thumb" that skewness or kurtosis divided by their respective standard errors are significant if less than −2 or greater than +2 to sketch the shape of the distribution curve.
9. can state in his own words, the concept of correlation between a dichotomous (two-valued) variable and a continuous (many-valued) variable as it relates to the item--total test score relation.
10. can find the best and worst discriminating item on a test using item analysis results.
11. can suggest reasons for an item's excellent or poor discriminating power.
12. can use item analysis information to edit test items he has written.

B. **PARTICIPANT'S PREPARATION:** Examine Computer Print-out Sheets

C. **INSTRUCTIONAL SESSION:** Discuss information to be gained from Computer Print-out Sheet

D. **PARTICIPANT'S PREPARATION:** Given the computer print-out sheet, complete the Study Questions for Analyzing Test Results.

E. **INSTRUCTIONAL SESSION:** Discuss participant's responses to Study Questions on Analyzing Test Results.

F. **MATERIALS:**

1. Computer Print-out Sheets
2. Study Questions for Analyzing Test Results
3. Test used to gather data for the computer print-out sheet.

G. **EVALUATION:** Proficiency Quiz No. 9

## ITEM DIFFICULTY

Item difficulty is defined as the proportion answering an item correctly. This is an unfortunate definition, because the harder the item the lower the proportion answering it correctly; i.e., the most difficult items have the lowest difficulty value.

There may be reasons for keeping or omitting easy or difficult items in a test. e.g., easy items provide motivation for the less able students and difficult items are a challenge to the more able students.

## ITEM DISCRIMINATING POWER

The correlation between correctly answering an item and receiving a high score on a test is called item discriminating power. A highly discriminating item will predict whether a subject falls into the high group or low group among test scores. Thus, one criterion for a test which will divide subjects into high and low groups is that the items be of sufficiently high discriminating power. (A typical minimum is .40)

One approximation to discriminating power is the difference of the proportions in the upper 27% and lower 27% correctly answering the item.

## CORRECTED SCORE

The corrected score is a means for adjusting a score for guessing. Frequently this is calculated by subtracting a fraction of the number wrong from the number correct. The fraction depends on the number of choices in the items. For example, if there are four choices per item, ¼ of the number wrong is subtracted from the number right. If the number of alternatives varies from item to item, the procedure is similar but the formula is more complicated.

SCORING AND ANALYSIS OF JR. HIGH SOCIAL STUDIES
INDIVIDUAL SCORES

| SUBJECT IDENTIFICATION | RAW SCORE | CORRECTED SCORE | OMITS |
|---|---|---|---|
| | 23 | 16 | 0 |
| | 26 | 20 | 0 |
| | 22 | 15 | 0 |
| | 22 | 15 | 0 |
| | 38 | 35 | 0 |
| | 38 | 35 | 0 |
| | 16 | 8 | 0 |
| | 39 | 36 | 0 |
| | 46 | 45 | 0 |
| | 22 | 15 | 0 |
| | 20 | 13 | 0 |
| | 25 | 19 | 0 |
| | 25 | 19 | 0 |
| | 19 | 11 | 0 |
| | 18 | 10 | 0 |
| | 44 | 43 | 0 |
| | 41 | 39 | 0 |
| | 42 | 40 | 0 |
| | 17 | 9 | 0 |
| | 20 | 13 | 0 |
| | 26 | 20 | 0 |
| | 42 | 40 | 1 |
| | 47 | 46 | 0 |
| | 10 | 0 | 0 |
| | 43 | 41 | 0 |
| | 19 | 11 | 0 |
| | 30 | 25 | 0 |
| | 20 | 13 | 0 |
| | 35 | 31 | 0 |
| | 46 | 45 | 0 |
| | 16 | 8 | 0 |
| | 37 | 34 | 0 |
| | 44 | 43 | 0 |
| | 44 | 43 | 0 |
| | 42 | 40 | 0 |
| | 39 | 36 | 0 |
| | 39 | 36 | 0 |
| | 25 | 19 | 0 |
| | 15 | 6 | 0 |
| | 41 | 39 | 0 |
| | 34 | 30 | 0 |
| | 45 | 44 | 0 |
| | 35 | 31 | 0 |
| | 37 | 34 | 0 |
| | 37 | 34 | 0 |
| | 41 | 39 | 0 |
| | 42 | 40 | 0 |
| | 39 | 36 | 0 |
| | 43 | 41 | 0 |
| | 32 | 28 | 0 |

## INDIVIDUAL SCORES

| SUBJECT IDENTIFICATION | RAW SCORE | CORRECTED SCORE | OMITS |
|---|---|---|---|
| | 30 | 25 | 0 |
| | 41 | 39 | 0 |
| | 40 | 38 | 0 |
| | 36 | 33 | 0 |
| | 36 | 33 | 0 |
| | 36 | 35 | 0 |
| | 33 | 29 | 0 |
| | 31 | 26 | 0 |
| | 17 | 9 | 0 |
| | 32 | 28 | 0 |
| | 25 | 19 | 0 |
| | 44 | 43 | 0 |
| | 26 | 20 | 0 |
| | 21 | 14 | 0 |
| | 37 | 34 | 1 |
| | 21 | 14 | 1 |
| | 34 | 30 | 0 |
| | 27 | 21 | 0 |
| | 37 | 34 | 0 |
| | 20 | 13 | 0 |
| | 34 | 30 | 0 |
| | 30 | 25 | 0 |
| | 36 | 33 | 0 |
| | 27 | 21 | 0 |
| | 39 | 36 | 0 |
| | 38 | 35 | 0 |
| | 24 | 18 | 0 |
| | 32 | 28 | 0 |
| | 31 | 26 | 0 |
| | 22 | 15 | 0 |
| | 37 | 34 | 0 |
| | 43 | 41 | 0 |
| | 22 | 15 | 0 |
| | 14 | 5 | 1 |
| | 34 | 30 | 0 |
| | 23 | 16 | 0 |
| | 36 | 33 | 0 |
| | 35 | 31 | 0 |
| | 24 | 18 | 0 |
| | 27 | 21 | 0 |
| | 26 | 20 | 0 |
| | 36 | 33 | 0 |
| | 33 | 29 | 0 |
| | 11 | 4 | 10 |
| | 34 | 30 | 0 |
| | 24 | 18 | 0 |
| | 16 | 8 | 1 |
| | 38 | 35 | 0 |
| | 24 | 18 | 0 |
| | 32 | 28 | 0 |

INDIVIDUAL SCORES

| SUBJECT IDENTIFICATION | RAW SCORE | CORRECTED SCORE | OMITS |
|---|---|---|---|
| | 41 | 39 | 0 |
| | 35 | 31 | 0 |
| | 42 | 40 | 0 |
| | 34 | 30 | 0 |
| | 36 | 33 | 0 |
| | 40 | 38 | 0 |
| | 41 | 39 | 0 |
| | 39 | 36 | 0 |
| | 38 | 35 | 0 |
| | 44 | 43 | 0 |
| | 32 | 28 | 0 |
| | 46 | 45 | 0 |
| | 39 | 36 | 0 |
| | 29 | 24 | 0 |
| | 34 | 30 | 0 |
| | 42 | 40 | 0 |
| | 38 | 35 | 0 |
| | 33 | 29 | 1 |
| | 39 | 36 | 0 |
| | 34 | 30 | 0 |
| | 44 | 43 | 0 |
| | 32 | 28 | 0 |
| | 33 | 29 | 0 |
| | 29 | 24 | 0 |
| | 35 | 31 | 0 |
| | 36 | 33 | 1 |
| | 39 | 36 | 0 |
| | 42 | 40 | 0 |
| | 41 | 39 | 0 |
| | 41 | 39 | 0 |

## ITEM INFORMATION

| ITEM NUMBER | DIFFICULTY | PROPORTION OF 130 SUBJECTS SELECTING EACH RESPONSE | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | OMIT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 0.846 | 0. | 0.846 | 0.069 | 0.031 | 0.054 | 0. | | | |
| 2 | 0.485 | 0. | 0.485 | 0.223 | 0.085 | 0.208 | 0. | | | |
| 3 | 0.423 | 0. | 0.123 | 0.423 | 0.277 | 0.169 | 0.008 | | | |
| 4 | 0.592 | 0. | 0.200 | 0.592 | 0.108 | 0.100 | 0. | | | |
| 5 | 0.923 | 0. | 0.023 | 0.008 | 0.038 | 0.923 | 0.008 | | | |
| 6 | 0.592 | 0. | 0.123 | 0.138 | 0.146 | 0.592 | 0. | | | |
| 7 | 0.915 | 0. | 0.062 | 0.915 | 0.015 | 0.008 | 0. | | | |
| 8 | 0.654 | 0. | 0.654 | 0.054 | 0.154 | 0.131 | 0.008 | | | |
| 9 | 0.585 | 0. | 0.031 | 0.585 | 0.362 | 0.023 | 0. | | | |
| 10 | 0.554 | 0. | 0.100 | 0.554 | 0.246 | 0.100 | 0. | | | |
| 11 | 0.615 | 0. | 0.277 | 0.615 | 0.077 | 0.031 | 0. | | | |
| 12 | 0.169 | 0.008 | 0.585 | 0.054 | 0.169 | 0.185 | 0. | | | |
| 13 | 0.662 | 0. | 0.177 | 0.146 | 0.662 | 0.015 | 0. | | | |
| 14 | 0.700 | 0. | 0.192 | 0.700 | 0.100 | 0.008 | 0. | | | |
| 15 | 0.608 | 0. | 0.608 | 0.123 | 0.185 | 0.085 | 0. | | | |
| 16 | 0.469 | 0. | 0.469 | 0.100 | 0.254 | 0.177 | 0. | | | |
| 17 | 0.869 | 0.008 | 0.015 | 0.062 | 0.869 | 0.046 | 0. | | | |
| 18 | 0.592 | 0.008 | 0.131 | 0.177 | 0.092 | 0.592 | 0. | | | |
| 19 | 0.600 | 0. | 0.600 | 0.115 | 0.069 | 0.215 | 0. | | | |
| 20 | 0.631 | 0. | 0.046 | 0.285 | 0.038 | 0.631 | 0. | | | |
| 21 | 0.669 | 0. | 0.162 | 0.669 | 0.062 | 0.108 | 0. | | | |
| 22 | 0.838 | 0. | 0.100 | 0.031 | 0.838 | 0.031 | 0. | | | |
| 23 | 0.623 | 0.015 | 0.138 | 0.123 | 0.623 | 0.100 | 0. | | | |
| 24 | 0.823 | 0. | 0.154 | 0.823 | 0. | 0.023 | 0. | | | |
| 25 | 0.669 | 0. | 0.062 | 0.092 | 0.177 | 0.669 | 0. | | | |
| 26 | 0.608 | 0. | 0.138 | 0.608 | 0.077 | 0.177 | 0. | | | |
| 27 | 0.254 | 0. | 0.331 | 0.254 | 0.246 | 0.169 | 0. | | | |
| 28 | 0.577 | 0. | 0.577 | 0.162 | 0.146 | 0.115 | 0. | | | |
| 29 | 0.485 | 0. | 0.100 | 0.485 | 0.138 | 0.269 | 0.008 | | | |
| 30 | 0.485 | 0. | 0.131 | 0.085 | 0.300 | 0.485 | 0. | | | |
| 31 | 0.615 | 0. | 0.192 | 0.146 | 0.615 | 0.046 | 0. | | | |
| 32 | 0.708 | 0. | 0.108 | 0.054 | 0.708 | 0.131 | 0. | | | |
| 33 | 0.615 | 0. | 0.615 | 0.108 | 0.085 | 0.192 | 0. | | | |
| 34 | 0.754 | 0. | 0.754 | 0.085 | 0.115 | 0.046 | 0. | | | |
| 35 | 0.515 | 0. | 0.515 | 0.154 | 0.238 | 0.085 | 0.008 | | | |
| 36 | 0.385 | 0. | 0.385 | 0.062 | 0.146 | 0.408 | 0. | | | |
| 37 | 0.877 | 0. | 0.046 | 0.877 | 0.062 | 0.015 | 0. | | | |
| 38 | 0.700 | 0. | 0.085 | 0.108 | 0.108 | 0.700 | 0. | | | |
| 39 | 0.492 | 0.008 | 0.492 | 0.138 | 0.169 | 0.192 | 0. | | | |
| 40 | 0.515 | 0. | 0.108 | 0.515 | 0.215 | 0.138 | 0.023 | | | |
| 41 | 0.946 | 0.008 | 0.015 | 0.015 | 0.015 | 0. | 0.946 | | | |
| 42 | 0.700 | 0.008 | 0.138 | 0.069 | 0.085 | 0.700 | 0. | | | |
| 43 | 0.846 | 0.008 | 0.054 | 0.846 | 0.069 | 0.023 | 0. | | | |
| 44 | 0.700 | 0.008 | 0.700 | 0.023 | 0.100 | 0.169 | 0. | | | |
| 45 | 0.738 | 0.008 | 0.138 | 0.031 | 0.738 | 0.069 | 0.015 | | | |
| 46 | 0.815 | 0.008 | 0.054 | 0.046 | 0.038 | 0.815 | 0.038 | | | |
| 47 | 0.846 | 0.008 | 0.846 | 0.038 | 0.069 | 0.015 | 0.023 | | | |
| 48 | 0.762 | 0.008 | 0.023 | 0.015 | 0.108 | 0.085 | 0.762 | | | |
| 49 | 0.854 | 0.008 | 0.069 | 0.854 | 0.023 | 0.031 | 0.015 | | | |
| 50 | 0.746 | 0.015 | 0.077 | 0.015 | 0.746 | 0.015 | 0.131 | | | |

TEST STATISTICS

| VARIABLE | MEAN | S.E. | ST. DEV. | S.E. | SKEWNESS | S.E. | KURTOSIS | S.E. | NO. OF SUBJECTS |
|---|---|---|---|---|---|---|---|---|---|
| RAW SCORE | 32.65 | 0.78 | 8.87 | 0.45 | -0.56 | 0.21 | -0.64 | 0.42 | 130 |
| CORRECTED SCORE | 28.34 | 0.97 | 11.03 | 0.55 | -0.54 | 0.21 | -0.69 | 0.42 | 130 |

SKEWNESS SIGNIFICANT AT .05 LEVEL

TEST RELIABILITY = 0.891
KUDER RICHARDSON FORMULA 20

ITEM ANALYSIS INFORMATION

| ITEM NUMBER | PROPORTION ANSWERING ITEM CORRECTLY | STANDARD DEVIATION OF ITEM | POINT BISERIAL CORRELATION OF ITEM WITH TOTAL TEST SCORE | POINT BISERIAL CORRELATION OF ITEM WITH CRITERION SCORE | RELIABILITY INDEX | VALIDITY INDEX |
|---|---|---|---|---|---|---|
| 1 | 0.846 | 0.361 | 0.337 | 0. | 0.121 | 0. |
| 2 | 0.485 | 0.500 | 0.336 | 0. | 0.168 | 0. |
| 3 | 0.423 | 0.494 | 0.296 | 0. | 0.146 | 0. |
| 4 | 0.592 | 0.491 | 0.412 | 0. | 0.203 | 0. |
| 5 | 0.923 | 0.266 | 0.280 | 0. | 0.075 | 0. |
| 6 | 0.592 | 0.491 | 0.191 | 0. | 0.094 | 0. |
| 7 | 0.915 | 0.278 | 0.302 | 0. | 0.084 | 0. |
| 8 | 0.654 | 0.476 | 0.086 | 0. | 0.041 | 0. |
| 9 | 0.585 | 0.493 | 0.315 | 0. | 0.155 | 0. |
| 10 | 0.554 | 0.497 | 0.267 | 0. | 0.133 | 0. |
| 11 | 0.615 | 0.487 | 0.114 | 0. | 0.056 | 0. |
| 12 | 0.169 | 0.375 | 0.097 | 0. | 0.036 | 0. |
| 13 | 0.662 | 0.473 | 0.156 | 0. | 0.074 | 0. |
| 14 | 0.700 | 0.458 | 0.387 | 0. | 0.177 | 0. |
| 15 | 0.608 | 0.488 | 0.471 | 0. | 0.230 | 0. |
| 16 | 0.469 | 0.499 | 0.221 | 0. | 0.110 | 0. |
| 17 | 0.869 | 0.337 | 0.393 | 0. | 0.132 | 0. |
| 18 | 0.592 | 0.491 | 0.251 | 0. | 0.123 | 0. |
| 19 | 0.600 | 0.490 | 0.243 | 0. | 0.119 | 0. |
| 20 | 0.631 | 0.483 | 0.253 | 0. | 0.122 | 0. |
| 21 | 0.669 | 0.470 | 0.538 | 0. | 0.253 | 0. |
| 22 | 0.838 | 0.368 | 0.332 | 0. | 0.122 | 0. |
| 23 | 0.623 | 0.485 | 0.486 | 0. | 0.236 | 0. |
| 24 | 0.823 | 0.382 | 0.443 | 0. | 0.169 | 0. |
| 25 | 0.669 | 0.470 | 0.419 | 0. | 0.197 | 0. |
| 26 | 0.608 | 0.488 | 0.421 | 0. | 0.205 | 0. |
| 27 | 0.254 | 0.435 | -0.147 | 0. | -0.064 | 0. |
| 28 | 0.577 | 0.494 | 0.340 | 0. | 0.168 | 0. |
| 29 | 0.485 | 0.500 | 0.477 | 0. | 0.238 | 0. |
| 30 | 0.485 | 0.500 | 0.416 | 0. | 0.208 | 0. |
| 31 | 0.615 | 0.487 | 0.297 | 0. | 0.144 | 0. |
| 32 | 0.708 | 0.455 | 0.483 | 0. | 0.220 | 0. |
| 33 | 0.615 | 0.487 | 0.481 | 0. | 0.234 | 0. |
| 34 | 0.754 | 0.431 | 0.552 | 0. | 0.238 | 0. |
| 35 | 0.515 | 0.500 | 0.557 | 0. | 0.278 | 0. |
| 36 | 0.385 | 0.487 | 0.113 | 0. | 0.055 | 0. |
| 37 | 0.877 | 0.329 | 0.362 | 0. | 0.119 | 0. |
| 38 | 0.700 | 0.458 | 0.522 | 0. | 0.239 | 0. |
| 39 | 0.492 | 0.500 | 0.355 | 0. | 0.177 | 0. |
| 40 | 0.515 | 0.500 | 0.321 | 0. | 0.160 | 0. |
| 41 | 0.946 | 0.226 | 0.251 | 0. | 0.057 | 0. |
| 42 | 0.700 | 0.458 | 0.482 | 0. | 0.221 | 0. |
| 43 | 0.846 | 0.361 | 0.546 | 0. | 0.197 | 0. |
| 44 | 0.700 | 0.458 | 0.498 | 0. | 0.228 | 0. |
| 45 | 0.738 | 0.439 | 0.482 | 0. | 0.212 | 0. |
| 46 | 0.815 | 0.388 | 0.496 | 0. | 0.193 | 0. |
| 47 | 0.846 | 0.361 | 0.572 | 0. | 0.206 | 0. |
| 48 | 0.762 | 0.426 | 0.451 | 0. | 0.192 | 0. |
| 49 | 0.854 | 0.353 | 0.316 | 0. | 0.112 | 0. |
| 50 | 0.746 | 0.435 | 0.534 | 0. | 0.232 | 0. |

NOTE. POINT BISERIAL CORRELATIONS OF ITEM WITH CRITERION SCORE AND VALIDITY INDEX ARE 0.0 BECAUSE OF ABSENCE OF CRITERION SCORE

## STUDY QUESTIONS FOR ANALYZING TEST RESULTS

### A. INTRODUCTION

The following set of questions is presented to help you to read the print-out which will accompany the corrected answer sheets returned to you after you pilot test items from the item bank. It may also aid you in making decisions concerning the results of your test.

While you are in the workshop, these questions may guide you in learning to interpret the print-out. Later, the questions may still be valid ones which you would ask yourself whenever you analyze test results, no matter in what special format the results appear.

### B. THE SET OF QUESTIONS

#### 1. QUESTIONS CONCERNING FORMAT AND GENERAL DATA

(1) In looking at the print-out sheets, are there variations in the format, that is, are the columns all in line, are there holes in the columns, etc.?

(2) Why did these variations occur? Did the students use the proper procedure for putting in ID numbers? Is your procedure for assigning ID numbers operational? That is, can you identify each student from the code?

(3) Are item numbers in consecutive order?

(4) Does the number of items reported correspond to the number of items on the test?

(5) Does the number of students correspond to the number of answer sheets you submitted?

(6) If a student number was omitted from the analysis, was the student absent that day?

(7) Are the items keyed correctly?

(8) In looking at the keyed answers, were the correct answers randomly placed in the a-e spaces, or was there, unknowingly, a pattern?

(9) What are the possible limits for each column on the output sheet?

(10) Is test reliability that is reported at an acceptable level?

(11) How do you determine if the distribution of scores that are reported represents a normal distribution?

(12) Is the distribution of scores on this test normal?

#### 2. QUESTIONS CONCERNING ITEMS ON THE TEST

(1) What proportion represents one person? (You use this in looking at the ITEM INFORMATION page of the print-out to see how many answered an incorrect alternative.)

(2) Are any "too easy" items listed? What are the limits which make an item too easy?

(3) Which is the easiest item?

(4) Why is this item the easiest?

(5) Do I want such an easy item on the test? If yes, Why?

(6) What is the point-biserial correlation of the easiest item and performance on the entire test?

(7) Are any "too difficult" items listed? What are the limits which make an item too difficult?

(8) Which is the hardest item?

(9) Why is this item the hardest? Is it keyed correctly?

(10) Do I want such a difficult item on the test? If yes, why?

(11) What is the point-biserial correlation of the hardest item and performance on the entire test?

(12) Do the more difficult items correspond to the higher level cognitive behaviors?

(13) Do the easier items correspond to the lower level cognitive behaviors?

3. QUESTIONS CONCERNING STUDENTS—INDIVIDUAL AND GROUP

(1) What is the range of raw and corrected scores?

(2) How is the corrected score derived?

(3) Which score do I want to record? Which score is more meaningful for me?

(4) What performance on this test will I accept as mastery of my objectives?

(5) How many got the easiest item wrong?

(6) Who got the easiest item wrong?

(7) How many got the hardest item wrong?

(8) Who got the hardest item wrong?

(9) Does the sign of the skewness and kurtosis correspond to my idea of the description of my classroom group?

4. QUESTIONS CONCERNING THE USES OF DATA FOR MAKING DECISIONS ABOUT INSTRUCTION

(1) How can I use all the information given in the data analysis as a guide for assigning grades?

(2) How can I use this information provided in the data analysis of my test to develop a re-teaching strategy?

(3) How can I change an easy item to make it more difficult?

(4) Do I want more easy items on the test?

(5) Did I think there were more easy items than the data shows? If yes, look for the items I thought were as easy as the easiest and find out why they were not as easy.

(6) Which alternative was chosen by those who got the easiest item wrong? What can we tell about their learning difficulties from their choice of the wrong alternative? Should I ask the ones who got the easiest item wrong to explain to me how they come to make this incorrect choice?

(7) How can I change a hard item to make it easier?

(8) Do I want more hard items on the test?

(9) Did I think there were more hard items on the test than the data shows? If yes, look for the items I thought were as hard as the hardest and find out why they were not as hard.

(10) Which alternative was chosen by those who got the hardest item wrong? What can I tell about their learning difficulties from their choice of the wrong alternative? Should I ask the ones who got the hardest item wrong to explain to me how they came to make this incorrect choice? Is their incorrect choice a better alternative, that is, is it a correct alternative also?

(11) How can I use the information given in the point-biserial correlation of an item with performance on the test?

(12) Which of the items are the best predictors of performance on the entire test?

(13) Could these best predictor items be used as the entire test?

(14) Could these best predictor items be used as a pre-test or a trial test?

(15) Why is a normal distribution important for me?

(16) If the distribution of scores is not normal, what does this mean?

(17) If the distribution of scores is not normal, is there another distribution with important characteristics which the sample of students which I have might fit better?

(18) If I have a non-normal distribution of scores, do I want to change the test so that for the same kind of students I will get a normal distribution? What must be done to accomplish this?

**Instructional Session Ten:** **Validity and Reliability**

**A. OBJECTIVES**

The participant

1. can match the four types of validity with their definitions.
2. can, by an illustrative example, express the degree of validity by a correlation coefficient.
3. can demonstrate knowledge of the term reliability by giving illustrative examples; e.g., repetition or equivalence.
4. can, by an illustrative example, express the degree of reliability by a correlation coefficient.
5. can, by an illustrative example, show that validity is limited by reliability.
6. can analyze a given testing situation by stating the relevant validity issue and the reason for the lack or presence of the type mentioned.
7. can, given an illustrative example, indicate the lack or presence of each of the four types of validity.

**B. PARTICIPANT'S PREPARATION:** Four Types of Validity

**C. INSTRUCTIONAL SESSION:**

1. Discuss the Four Types of Validity.
2. Completion of the Practice Sheet on Validity by participants using Four Types of Validity as source material.
3. Group critique of participant's responses on Practice Sheet.

**D. PARTICIPANT'S PREPARATION:** Reliability of Test Scores.

**E INSTRUCTIONAL SESSION:**

1. Discuss the Reliability of Test Scores.
2. Discuss the relationship between reliability and validity.

**F. MATERIALS:**

1. Four Types of Validity
2. Validity - Practice Sheet
3. Reliability of Test Scores

**G. EVALUATION:** Proficiency Quiz No. 10

## FOUR TYPES OF VALIDITY

### CONTENT VALIDITY

– the degree to which a test represents behaviors and subject matter of the course objectives

VALID

a test constructed using a "behavior by content" table of specifications based on the course objectives

INVALID

a teacher emphasizes "critical thinking" in Social Studies class and gives a test requiring only recall of specific facts

### CONCURRENT VALIDITY

– the degree to which a test reflects a separate simultaneous judgment

VALID

scores on a teacher-made test in science correlate highly with scores on a standardized test given the same week

INVALID

scores on a test are very different from what the teacher had expected on the basis of her experience with the class

### PREDICTIVE VALIDITY

– the degree to which a test is able to predict some specific future outcome

VALID

College Board scores correlate highly with collegiate academic performance

INVALID

no prediction possible

### CONSTRUCT VALIDITY

– the degree to which a test correctly judges a psychological concept or theory

VALID

test items that measure higher mental processes have higher difficulty level

INVALID

no correspondence to a psychological theory

## VALIDITY – PRACTICE SHEET

A. What type of validity is primarily involved in each of the following situations?

1. A fourth grade teacher prepares a table of specifications to compare the items of her test with the objectives of the unit.

   a. content
   b. concurrent
   c. predictive
   d. construct

2. The U. S. Employment Service wishes to test men to determine which ones have had enough experience to be referred to contractors who have vacancies for electricians.

   a. content
   b. concurrent
   c. predictive
   d. construct

3. A medical school wishes to test the personalities of its applicants to determine which ones are best suited to a physician's responsibilities.

   a. content
   b. concurrent
   c. predictive
   d. construct

4. A pencil-paper test is used to identify students entering junior high school who have emotional difficulties and should be singled out for counseling.

   a. content
   b. concurrent
   c. predictive
   d. construct

5. The Short Employment Tests are found to correlate .91 with the General Clerical Test, which has been used for some time as a predictor of job success.

   a. content
   b. concurrent
   c. predictive
   d. construct

B The State of Illinois administers written tests and actual performance tests to driver license applicants.

1. How would you go about determining the content validity of these tests?

2. What assumption or assumptions would have to be made before these tests could be used to predict the licensee's actual performance on the highways?

## RELIABILITY OF TEST SCORES

I. **DEFINITION OF RELIABILITY:** Reliability refers to the **consistency** of test scores, i.e., how consistent a set of test scores measures whatever *it does* measure.

II. **METHODS OF IMPROVING TEST RELIABILITY:**

A. **Increasing** the length of the test yields an increased number of samples of pupil behavior. A larger sample usually increases the spread or range of the scores.

B. **Insuring** that the items on the test are **homogeneous,** i.e., they measure the same behavior.

C. **Increasing the discrimination** power of the items, i.e., assist the teacher in identifying the better and the poorer students.

D. **Utilizing** items that are of **middle difficulty** rather than only items that are very difficult or very easy.

E. **Administering** the test to a group having a **wide range** of **ability** levels.

III. **METHODS OF OBTAINING VALID DATA TO COMPUTE RELIABILITY COEFFICIENTS**

A. **Test - Retest Method:** Administering the same test twice to the same class.

1. Advantages: Only one test is needed.

2. Limitations:

a. Student may recall items from the first test situation.

b. Student may have "learned" material related to the test in the interval between test and re-test.

c. Classroom time may be wasted by the repeated administering of the test.

d. Student often loses interest in the second testing situation.

B. **Equivalent Forms:** Parallel forms of a test are administered during the same test situation.

1. Advantages: Reduction of limitations "a" and "b" of Test-Retest Method.

2. Limitations:

a. Educational achievement tests are usually not published in alternative forms.

b. Behavior sampled by the two tests may not be identical.

C. **Split-Halves Method:** Single test is split into two equivalent halves. Common method of doing so is to score the odd-numbered items and the even-numbered items separately.

D. **Kuder–Richardson:**

Formula $$r = \frac{k}{k-1}\ \frac{[\,1 - \Sigma pq\,]}{\sigma^2}$$

k= number of items in the test

$\Sigma$ = sum

$p$ = proportion of the population answering one item correctly

$q$ = proportion of the population answering one item incorrectly

$\sigma^2$ = variance of the total scores on the test

## Instructional Session Eleven: The Liaison Role of the Workshop Participant

A. **OBJECTIVE**

The participant

1. can discuss the role of the Workshop Teacher as an agent of change.
2. will continue an active relationship with the Evaluation Project.
3. will share ideas on applying the knowledge and skills learned with other participants and other members of the school district.

B. **INSTRUCTIONAL SESSION:** Review and discuss materials on "The Workshop Teacher as Liaison."

C. **MATERIALS:**

1. The Workshop Teacher as Liaison.

87

## THE WORKSHOP TEACHER AS LIAISON

After returning to his school, the workshop teacher will probably be asked to function as a "liaison" or link between the Evaluation Project and his district mainly for the purpose of sharing ideas about the objectives of the workshop and how they relate to the objectives in his particular area of specialization and the general curriculum of the school district. It is our hope that existing evaluatory practices will be augmented to some degree through proper utilization of the workshop teacher and available resource persons from the Evaluation Project. This liaison role has several important aspects and the emphasis given to each will vary from district to district according to its needs. Sub-roles which the teacher may assume as "Liaison" are:

1. **EVALUATOR** - As an evaluator, he will be able to

   a. **recognize** the important relationship between the objectives of instruction and the construction and uses of tests as they affect improved teaching and learning.

   b. **distinguish between** effective and ineffective teaching-testing practices.

   c. **select** and/or **develop** more suitable test instruments to evaluate the individual student's progress.

2. **DISSEMINATOR** - As a disseminator, he will be able to

   a. **share** information with his school administration and other classroom teachers regarding his experiences as a workshop participant.

   b. **assist** local administrators and teachers who are concerned with the problem of planning a school-wide testing program.

   c. **discuss** the appropriateness of his experience and plans for implementation within their classroom structure.

   d. **develop** plans for in-service workshops in evaluation with assistance from the Evaluation Project staff.

   e. **distribute** the Operational Guide and other materials produced by the Evaluation Project staff to interested teachers for purposes of preview.

   f. Individual districts may request that the participant submit a written report to his Principal, Superintendent, or Curriculum Specialist. Emphasis should be placed on the objectives of the Evaluation Project workshop; the value of the experience; an orderly account of **all** activities pursued during the interim of the workshop; suggestions for improving "future" evaluation workshops; and plans for following activities including implementation of new instructional and evaluative practices within his area of competency.

3. **LIAISON** - As a liaison person, he will be able to

   a. **utilize** the services of the Evaluation staff for assistance in conducting in-service workshops.

   b. **request** that his name be placed on the mailing list for the purpose of receiving information on new project activities and materials.

   c. **attend** the continuing contact meetings which are held twice yearly.

## THE ITEM POOL — Compilation, Testing, and Projected Use

A tangible outcome of the Evaluation Project will be the creation of an item pool of multiple choice items matched with behavioral objectives. This bank of test questions will be available to all districts who are participants in the Evaluation Project. The purpose of the Project item bank is to provide a source of reliable test questions from which teachers can select those appropriate to their on-going classroom instruction at any point throughout the school year.

### Compilation Of The Item Bank

In a series of workshops, classroom teachers write instructional objectives and test items to measure attainment of those objectives. After items are written, they are subjected to at least four different edits. The writer edits them first and in so doing, improves his technique of writing well-formed items. A member of the Project staff then edits the items for form and discusses them with the teacher. The items are then given to a subject matter specialist who does the third editing. A professional editor will complete the final editing.

### Testing Items In The Bank

As part of the editing process items will be pilot tested on a continuing basis. Only when the items have been shown to be statistically reliable and valid will they go into the final test item pool for mass distribution. Even then, if necessary, they will continue to be edited.

If possible, the first pilot testing of an item will be done by the writer of the item. The revised item will then be sent for as many more pilot tests as statistical studies of it indicate are necessary. With each testing, the writer is furnished an analysis of the test items. (See Sample) On the basis of this analysis, the writer may alter items, discard them, and write additional items. As items are pilot tested, difficulty level is constantly up-dated.

### Projected Use Of The Item Bank

When the item pool becomes operational, the objectives will be coded in a multi-category system. They will be coded by behavior and by content. The behavior may be categorized as recall or above, or the code may be based on a finer categorization of 1-6 based on Bloom's Taxonomy. Possibly, a combination of these two methods may be adopted. The content will be distributed among four general categories of mathematics, science, language arts, and social studies. A finer sub-dividing of each of these four categories will be made based on the exhaustive content listings compiled by the teachers and on grade level.

Besides the data indicated by the coding process, the difficulty level will be furnished for each item selected by a teacher.

Policy and practice on item retrieval is still being established. It can work in several ways. Teachers might be given a book containing all coded objectives under one content area at a particular grade level. Numerous test items would be listed under each objective. The teacher can then select specific test items to make up a test. Another method would be to order a test by specifying the percent of items at each level of behavior and difficulty level for a particular unit of content, for example, the Civil War. The test would then contain a random sample of items from the pool which meets those specifications. A third use of the item pool would be for diagnostic testing. Because each objective has more than one item under it, a teacher could order a pre- and post-test, being assured that items in each test are of comparable difficulty and test the same behavior and content.

## Reference Materials


1. Bloom, Benjamin S. (editor), Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook 1: Cognitive Domain, New York: David McKay Co., Inc., 1956.

2. Lennon, Roger T., "A Glossary of 100 Measurement Terms," Test Service Notebook No. 13, Issued by Test Department, Harcourt, Brace and World, Inc., New York.

3. Lindquist, E. F. and A. M. Hieronymous, Iowa Test of Basic Skills and Teacher's Manual, Prepared at the State University of Iowa, Grades 3 - 9, Forms 1, 2, 3 and 4, Houghton Mifflin Co., New York, 1964.

4. Lindvall, C. M., Testing and Evaluation: An Introduction, Chicago: Harcourt, Brace and World Inc., 1961.

5. Popham, W. James, Educational Statistics: Use and Interpretation, Evanston: Harper and Row, 1967.

6. Mager, Robert F., Preparing Instructional Objectives, Palo Alto, California: Fearon Publishers, 1962.

7. Talmage, Harriet, "Unit Two: Operationalizing Instructional Objectives," Improving Teaching Skills, Chicago: Science Research Associates, Inc.,

8. Test Development Division of Educational Testing Service, "Multiple-Choice Questions: A Close Look," Princeton, New Jersey: Educational Testing Service, 1963.

9. Tyler, Ralph W., Basic Principles of Curriculum and Instruction, University of Chicago, Chicago, Illinois, 1968.